UNIVERSITY OF DELHI
SCIENTIA ET MORES

RATAN TATA
LIBRARY

DELHI SCHOOL OF ECONOMICS

THE RATAN TATA LIBRARY

Cl. No. B280X H7

Ac. No. 12630

Date of release for loan

This book should be returned on or before the date last stamped below. An overdue charge of twelve nP. will be levied for each day/hour the book is kept beyond that date.

~~MAR 1963~~
~~13 MAR 1963~~
~~12 NOV~~ 1975
~~28 OCT 1963~~
~~6 NOV 1963~~
~~2 NOV 1963~~
~~18 NOV~~ 1963
25 NOV 1963
9 DEC 1963
27 APR 1964
2 ~~5~~ MAY 1964
2 [illegible] SEP 1967

A COURSE IN
BUSINESS STATISTICS

PUBLISHED BY PITMAN

STATISTICS IN THEORY AND PRACTICE

By L. R. CONNOR, M.Sc., *Barrister-at-Law*

Part I. STATISTICAL METHOD. Introduction—Organization of a Statistical Inquiry—Statistical Data—Statistical Measurement—Classification and Tabulation—Diagrams—Graphs—Derivation Data—Statistical Groups—Statistical Averages—Dispersion—Skewness—Probability and Error—Sampling—Correlation—Index Numbers—Finite Differences, Interpolation, Graduation, and Least Squares.

Part II. APPLIED STATISTICS. Introduction—Population—Prices—Wages—Employment—Profits—Trade—Finance—Production—Wealth—Business Barometers and Business Activity Indices—Business Statistics—Appendix II: Reference to Part II—Index.

In medium 8vo, cloth gilt. 12*s*. 6*d*. net.

BUSINESS CHARTS

By T. G. ROSE, M.I.Mech.E.

With a Foreword by SIR FRANCIS W. GOODENOUGH, C.B.E.

Value of Graphics—Illustration Charts—Circle or Pie Charts—Symbolic Charts—General Rules for Graphics—Bar Chart—Line Chart—Z Chart—Rate of Change Charts—Multiple Control Charts—Index.

In demy 8vo, cloth gilt, 104 pp. 7*s*. 6*d*. net.

STATISTICS AND THEIR APPLICATION TO COMMERCE

By A. LESTER BODDINGTON

Fellow of the Royal Statistical and Economic Societies
Member of the American Statistical Association

Part I. Business Statistics and their Importance—Compilation and Uses of Statistics—Accuracy and Approximation—Types and Averages.

Part II. Tabulation—Graphic Method—Graphic Method Applied to Commerce—Methods of Comparison—Index Numbers—Correlation—Census as Aid to Scientific Business—Business Research—Mechanical Aids to Statistical Work—Tables—Index.

In medium 8vo, cloth gilt, 351 pp. 12*s*. 6*d*. net.

A COURSE IN BUSINESS STATISTICS

THE ELEMENTS OF STATISTICAL METHODS WITH EXERCISES AND ANSWERS

BY

CHARLES H. BAYLISS
B.Com.(Lond.)

LONDON
SIR ISAAC PITMAN & SONS, LTD.
1947

SIR ISAAC PITMAN & SONS, Ltd.
PITMAN HOUSE, PARKER STREET, KINGSWAY, LONDON, W.C.2
THE PITMAN PRESS, BATH
PITMAN HOUSE, LITTLE COLLINS STREET, MELBOURNE
UNITEERS BUILDING, RIVER VALLEY ROAD, SINGAPORE
27 BECKETTS BUILDINGS, PRESIDENT STREET, JOHANNESBURG

ASSOCIATED COMPANIES
PITMAN PUBLISHING CORPORATION
2 WEST 45TH STREET, NEW YORK
205 WEST MONROE STREET, CHICAGO

SIR ISAAC PITMAN & SONS (CANADA), Ltd.
(INCORPORATING THE COMMERCIAL TEXT BOOK COMPANY)
PITMAN HOUSE, 381–383 CHURCH STREET, TORONTO

BOOK
PRODUCTION
WAR ECONOMY
STANDARD

THE PAPER AND BINDING OF
THIS BOOK CONFORM TO THE
AUTHORIZED ECONOMY STANDARDS

MADE IN GREAT BRITAIN AT THE PITMAN PRESS, BATH
D7—(B.38)

PREFACE

In compiling this book on Business Statistics I have given particular attention to the requirements of students desiring a knowledge of statistics.

An endeavour has, therefore, been made to state clearly and simply the fundamental principles of the Theory of Statistics, and to illustrate them by examples drawn from experience of the practical application of these principles to commercial problems of various kinds.

It is hoped that the book will prove of value to those to whom the study of statistics is essential for examination purposes, and also of assistance to those who may be called upon to apply statistical methods in their business.

My thanks are due to Mr. D. L. C. Rees for valuable suggestions and useful criticism.

C. H. B.

CONTENTS

CHAP. PAGE

PREFACE v

I. OBJECTS AND IMPORTANCE OF STATISTICS . 1

II. COMPILING DATA 8

III. ACCURACY AND APPROXIMATION . . . 26

IV. THE ARITHMETIC AVERAGE. . . . 38

V. THE MOVING AVERAGE 56

VI. THE MEDIAN AND THE MODE . . . 66

VII. INDEX NUMBERS 78

VIII. TABULATION 91

IX. THE GRAPHIC METHOD 102

X. DIAGRAMS 117

XI. CORRELATION 126

XII. BUDGETARY CONTROL IN BUSINESS . . 138

XIII. PUBLISHED STATISTICS 145

ANSWERS TO THE EXERCISES . . . 155

INDEX 175

A COURSE IN BUSINESS STATISTICS

CHAPTER I

OBJECTS AND IMPORTANCE OF STATISTICS

THE ever-increasing complexities and the keener competition of modern business have created the need for improved methods of control. Previously the success of an enterprise depended largely on the wisdom and foresight of its leaders, and the measure of success it obtained was the result of the accuracy of this foresight. In so far as men failed to realize the trend of events with only individual opinion to guide them, so they failed to remain in business. The tendency for industry to be carried on more and more by large organizations has caused this method of control to become impossible, while the greater difficulties encountered in working profitably have made it imperative that the direction of commercial policy should be based on facts, the accuracy of which is essential. It is, therefore, due to this revolutionary change in business organization that more scientific methods have been devised, and that the importance of statistics, as providing for this, is realized to an unprecedented extent.

The problem of making increasing profits in the more difficult conditions which now exist resolves itself into the necessity for increasing turnover, and reducing manufacturing costs by securing the utmost efficiency in production. This can only be achieved by continually having a finger on the pulse of the business, so that its progress is always under observation, and, in addition, securing as complete a realization as possible of the economic conditions governing the supply of raw materials and the markets for the finished products. Decisions must be based on facts, showing the tendency of sales to increase or decrease, under the influence of fashions, competition, or substituted products. The comparison between the price obtained over various periods, the trend of expenses, and every factor influencing costs of production must receive due consideration. The guiding policy of a business must, therefore, rest upon market research, statistics of prices, and data concerning commerce and industry in general. It is only by continuous study of facts such as these, relating to the past and the present, that reliable estimates can be obtained for the future. Only by such scientific methods is it possible for an organization to be kept at the peak of competitive efficiency, and its control to be based on a comprehensive conception of the factors determining its success. It is because this position has been realized that many organizations have found it profitable to establish statistical departments whose function it is to amalgamate and co-ordinate the data relating to the individual business, to carry out market research, and, in addition, to investigate all the relevant facts concerning commerce as a whole.

The Objects of Statistics

Statistics, therefore, provide the means whereby business may be controlled in the most efficient manner, and it is now possible to summarize the objects of statistics as follows—

1. To present facts in such a manner as to show precisely what has occurred in the past. By analysis and comparison to show clearly such changes as have taken place, and to investigate the whole of these facts in relation to all available evidence relevant to the questions under review.

2. To obtain data relating to the present to provide comparisons showing whether conditions operating in the past are being maintained or improved upon, so that constant control is retained over important matters.

3. To utilize the data so obtained in such a manner as will enable reliable estimates of future conditions to be made. In this way decisions may be based on facts, and on estimates which are likely to prove reasonably correct.

The objects outlined above are applicable equally to the internal management of an industrial concern, its relations with similar organizations, and the direction of its general policy.

Statistics may, therefore, be defined as the presentation of facts numerically in a manner designed to enable the relationships between the various facts to be ascertained.

Statistical Methods

In the application of these principles of statistics to the actual control of business, the presentation of data must always be given serious consideration.

Frequently figures are produced which either through lack of knowledge or unscrupulousness are entirely misleading, and it is to ensure that correct conclusions may be deduced from the data produced that statistical methods have been developed.

In the interpretation of statistical data it is, however, imperative to remember that in the investigation of any problem there are frequently facts which, although having a material bearing on the question, cannot be expressed as numerical statements. Such details as these should not be overlooked in formulating decisions on the information statistically presented.

Finally, a correct knowledge of the statistical methods adopted is also essential. Types such as averages, may otherwise be employed in such a manner as to lead to erroneous conclusions solely due to the fact that the limitations of the types employed are not sufficiently appreciated.

The following example will show how vitally important it is to consider the methods to be used for the presentation of data.

A limited company is being formed to take over the business of two existing firms controlled by an individual proprietor. The promoters of the new company desire statistics of the turnover and profits of the two businesses which they propose to purchase, but the vendor will not disclose actual figures. Instead he assures the promoters that the turnover of the two businesses has increased by 10 per cent, business X showing an increase of 20 per cent, and business Y a decrease of 10 per cent. This statement, however, creates a false impression, because the methods adopted in arriving at the net increase of 10 per cent are incorrect. This is shown by the following table, which discloses the actual figures relating to the two concerns, from which it will be seen that owing to the difference in magnitude of the figures under consideration there is actually a net decrease in the turnover.

	First Year	Second Year	
	£	£	
Business X. .	10,000	12,000	– 20% increase
Business Y . .	100,000	90,000	– 10% decrease
	110,000	102,000	– Actual decrease $7\frac{3}{11}$ %

As a further example, the comparisons usually made when studying Trading and Profit and Loss Accounts may be considered. Immediately comparisons are made between the accounts for one period and those for another, statistical methods are employed. In doing this it may be found that the profits have increased from £100,000 to £120,000—on the face of it a highly satisfactory result. But this conclusion ought not to be made without first taking into account other factors. For example, has the profit been made on the same amount of capital, since it is reasonable to expect that increased capital will result in increased profits? Or, again, have the profits increased proportionately to the sales? If the sales have increased by a higher percentage than the profits, it is obvious that the cost of obtaining these sales has increased disproportionately, as a result of increased manufacturing costs, or of increased management expenses, or of higher selling costs. Investigations in these directions may reveal that intensive advertising or other sales charges have resulted in higher sales, but that the actual increase in profit does not justify the expenditure. On the other hand, it may be found that while the turnover may be equal to or greater than that for the previous year, the actual quantity of goods produced and sold has increased far more, indicating that the prices obtained for the products of the company have been reduced.

The above very clearly illustrates the erroneous conclusions which are likely to result when due consideration is not given to the appropriateness or the statistical accuracy of the methods employed.

The Value of the Applications of Statistics

The necessity for the use of statistics in the control of business has been sufficiently emphasized, but an example which shows directly the advantage to be derived from the use of statistical methods may usefully be considered at this point.

The proprietor of a business was recently considering the results of his trading, and, although the profits were reasonably satisfactory, he decided to investigate statistically the contributions made to the general result by each independent department. Although the costs indicated that one department should have made a considerable profit, his investigations disclosed the fact that this department was making a loss. This immediately led to further investigations, which ultimately revealed that, owing to the fraudulent practices of certain employees, he was losing considerable amounts of money. In this manner, a loss was subsequently converted into a profit, thereby improving the results of the concern as a whole.

The illustration which has been considered is small compared with the extensive applications which it is possible to make in large organizations. In studying various statistical methods, reference will be made to the many problems which may be treated statistically, thereby assisting in attaining that maximum efficiency in production and marketing which is so vital to the success of a commercial enterprise.

EXERCISE

1. What is the difference between statistics and mere figures?

2. Why is it frequently said that statistics can be made to prove anything, and why is this statement untrue?

will provide information to show the trend of expenses and standing charges, the effect of sales policy and the degree of efficiency which is being maintained.

For the efficient control of a manufacturing concern a considerable amount of data will be required concerning production, and the collection of this data will call for careful organization. Each business will have its own particular problems. In almost all cases special attention will require to be given to the output of machines in relation to their capacity. It is only through the investigation of all causes resulting in the output of any machine falling below a previously fixed level that the utmost efficiency can be obtained. Statistics must also be available concerning the output of individual workers, to enable methods of payment to be decided upon, to secure the conditions under which the maximum output is obtained, and to reduce fatigue to a minimum. Wastage in all its forms must be controlled by means of statistics. Probably a most important form of wastage—that due to trained employees leaving—is frequently overlooked, although in many cases it is a serious problem which, if statistically analysed and investigated, could be avoided.

It has already been seen that statistics of trade in general are also important in the control of business, and there has grown up a considerable number of sources from which this information is available. The methods of compiling this data are numerous, but may be illustrated by a consideration of the import and export statistics of the United Kingdom. Every exporter is compelled by law to send to the appropriate customs officer a statement of the quantity and value of the goods he exports, together with a description of the merchandise. To ensure that all exports are included, the whole of the cargo of every ship has to be accounted for before the ship is allowed to leave port. Imports have to pass through the customs for the purpose of assessing duty, but also in order that records may be kept of all goods entering the country. When the data so obtained are classified and analysed, valuable information is available concerning imports and exports, indicating the amount of foreign competition which each trade has to meet, the extent to which foreign markets are being held, and the volume and variety of foreign trade as a whole.

In similar ways data are compiled showing the activity of various industries, the production of coal, pig iron and steel, the movements of price levels, the quantity of goods carried by the railways and many other factors which are valuable barometers of the trend of business activity.

The Questionnaire

In many investigations the questionnaire is employed to obtain the necessary data. The problems to be investigated having been carefully considered, it is frequently necessary to obtain information from a large number of sources. This is generally accomplished by issuing to all those interested or from whom information is desired a form of inquiry, known as a questionnaire. Probably the most well-known questionnaire is the one issued to every householder for the purpose of compiling the census of population. The same method is employed in compiling the census of production, in which case information is demanded from every manufacturer. These are, however, not the only instances in which the questionnaire is employed, and its use in commercial investigations is increasing. For example, the cost of carriage enters very largely into the cost of goods purchased by a large undertaking, and, in many cases, investigations have been carried out with the object of reducing the cost to a minimum. To do this questionnaires are sent to all suppliers to obtain complete data as to their methods of consigning goods. The results are then tabulated and the outcome of a study of the varying methods employed is frequently a standardization of procedure resulting in extensive savings on carriage charges, ultimately benefiting the purchaser and the vendor.

Under the present organization of industry it is frequently necessary for all manufacturers to combine in common action designed to meet the needs of the entire industry. In the majority of such cases it is usually important to obtain information relevant to the particular problem from all those concerned, and this information is gathered by sending questionnaires for completion to all those interested.

It will be seen that if the method is to achieve its object, great care must be taken in drawing up the questionnaires. As far as possible all questions should call for a simple answer; very careful consideration should be given to the class of people who are to complete them, so that the inquiry will not be rendered useless by the questions being so designed that they are difficult to answer definitely. In addition to this, it should be borne in mind that when the answers are received the results have to be tabulated, and a clear conception of the tables ultimately to be compiled should be formed before the questionnaire is drafted. The following is a typical example of a questionnaire—

QUESTIONNAIRE

Name of firm	
Address	
Occupation	
Number of oxy-acetylene plants in operation . . .	
How long has oxy-acetylene welding been carried on? .	
Average quantity of acetylene used per week . . .	
Are any extensions to the present plant contemplated? If so, give some indication of the magnitude of the alterations .	
Number of employees engaged on oxy-acetylene welding .	

Sampling

When the census of population is undertaken, it is obvious that to secure data likely to be of any value the greatest care must be taken to cover as far as possible every inhabitant of the country. To achieve this purpose it is specifically laid down in the questionnaire who is to be included in the return and the exact time to which the return relates. As far as possible, no dwelling place is omitted, and every endeavour is made to ensure that the utmost accuracy is obtained, even to including those of no fixed abode. Again, in compiling details of export and import trade, every precaution is taken to include every item of trade and (theoretically at least) nothing is omitted, although in actual practice this may not be so.

The rigorous attention to detailed accuracy employed in these statistics is, however, not possible in every inquiry. In some investigations it is impossible to obtain exact figures, while in others the cost of carrying out the investigation in absolute detail would not be justified by the ultimate usefulness of any results obtained. This difficulty obtains occasionally when the problem concerns only a single business, but most frequently it applies to questions relating to a particular trade or to the commerce of a whole country.

Attempts to deal with such extensive problems have resulted in recent years in the development of a statistical method known as the "Method of Sampling," which is largely based on the Law of Statistical Regularity. To illustrate this law, suppose that a penny is tossed into the air, then nothing in our knowledge enables us to decide whether it will come down heads or tails. But if a million tons of pennies are thrown, we know that 500,000 tons will come down heads and 500,000 tons tails. It may not be possible to toss this quantity of pennies, but if we start with throwing a few pennies, and gradually increase the number thrown, we shall find that the more throws we make, the nearer we shall approximate to equal quantities of heads and tails. If a great number of throws are made, the result will differ from equality by only a small percentage, which will decrease as the quantity thrown increases.

The Law of Statistical Regularity, therefore, states that if a considerable number of examples are taken, these examples will be fairly representative of the whole from which they are taken. The accuracy of this statement may be proved by taking the number of letters on any ten lines on this page, and obtaining the average number of letters per line. If this average is then multiplied by the total number of lines on the page, taking into account short lines, it will be found that the estimated number of letters on the page so obtained will be in close proximity to the actual number of letters if they are accurately counted.

The application of this principle to commercial investigation can, therefore, be expressed in the fact that when it is impossible to consider a mass of detail, if a sufficiently large number of samples are taken at random, they then may be assumed to be representative of the whole. In virtue of this, the method can be extensively applied to problems which it would otherwise be impossible to investigate. Valuable information concerning the profits of industry, for example, can be obtained in this way. It is obviously impossible to obtain the trading results of every company carrying on business, partly because the information is not available concerning many enterprises privately controlled, or organized as Private Limited Companies, and, in addition to this, it would be extremely difficult to tabulate the results in such a manner that the desired information would be satisfactorily obtained. If, however, it is possible to obtain a reasonable number of representative reports, it can be satisfactorily concluded that the result is fairly representative of industry as a whole. If the profits of, say, 2,000 companies, chosen at random, and including as many industries as possible, are taken, and it is found that the profits of these companies represent a return of 5 per cent on the invested capital, the result will be very little different from the results which would have been revealed had it been possible to obtain the complete data. At the same time should it be found that the profits for one year for the companies actually considered show a decrease as compared with the corresponding figures for the previous year, then it will be reasonable to conclude that this decrease is significant of the trend of profits for all industry.

Similar methods may be employed for many investigations of this nature. When it is desired to obtain statistics to show the manner in which the average man spends his income, it is obviously impossible to investigate every case. A sufficiently accurate indication can, however, be obtained by sampling, choosing examples at random unbiased by any preconceived idea of the results of the investigation. Overcrowding, earnings of workmen, productivity of agricultural land, and similar problems can be solved by this method.

Sampling is also frequently valuable in investigations relating to the control of an individual business.

A typical example may be cited in connection with employees' earnings. When it is desired to fix a piecework price for a particular operation, it is essential to know the output it is possible for a worker of fair average capacity to obtain. It is upon the output of such a worker that a reasonably accurate price can be fixed, in relation to a preconceived basic wage. When a large number of workers are engaged on the same work, it is, however, impossible to record the capacity of every worker, but if a fair proportion of the workers are investigated, the results obtained will give a satisfactory average on which to work. In such an inquiry, care must be exercised to ensure that representatives of all grades of workers on the operation are included. It is quite possible that by reason of unconscious bias, only the best workers may be chosen, and to avoid the possibility of erroneous conclusions such as would result from anything of this nature, it is a frequent practice to choose the samples by lottery.

Market research has become increasingly important as a means of obtaining greater sales, and in fact the striking successes which have attended the marketing of certain proprietory lines were based primarily on the thorough market research that preceded the initiation of the sales campaign. The sales of similar lines by competitors are frequently far less, although in point of price and quality they compare favourably. This position largely results from failure to plan the sales and advertising policy on an accurate knowledge of the markets. Yet this knowledge of the markets can be obtained by sampling. The answer to a question such as: "What brand of our product has made the greatest sales headway in the past year, and why?" may be the solution to many serious troubles. An investigation among dealers, representative of all those selling the product, in middle, lower middle, and working class districts, may provide data which, although not comprehensive in that every dealer is included, will be so representative of the whole that the information may be taken as giving an accurate estimate of prejudices of wholesalers, consumers' preferences, and the effects of advertising policy.

If it is found that sales in any particular area are falling, by adopting the method of sampling and investigating the facts among representative traders, the cause can be ascertained and steps taken to regain the lost turnover.

Characteristics of Statistical Data

By whatever method statistical data is compiled it is important that the data conform to certain characteristics, and in interpreting the results of any inquiry, consideration must first be given to the extent to which these requirements are satisfied. Thus, data must possess the following attributes: (1) Comparability, (2) homogeneity, (3) completeness, (4) stability, (5) representativeness, (6) freedom from personal bias. Frequently statements are made which are based on statistics, but, owing to the fact that sufficient attention is not given to these characteristics, the conclusions are misleading. In compiling some data it is impossible to fully satisfy the requirements mentioned above, and, in such cases, it is important that the extent to which this is impossible should be emphasized.

Comparability

Statistics are largely concerned with making comparisons, but comparability will only be possible when the unit for the data conforms in all cases to a definite standard. Having, therefore, decided on the scope of any inquiry, and planned the data to be collected in order to obtain a solution of the problem, attention must then be given to the unit in which this data has to be expressed. The unit must be such that while it remains constant itself, it will be capable of including any variable factors which have an effect upon it. An example will make this seeming paradox clear.

A works manager desires to make a comparison between the present output of his factory and the output for a similar period in a previous year. He may compare the output for a week with the actual output for a week in the previous period, thus making the weekly output of the factory his unit. In doing so he may find that his output has increased considerably and, therefore, assume that his operations have been successful. This, however, may not be wholly true, because his unit of data has ignored two important factors, each having a bearing on his output—time and labour. His comparison is influenced by overtime or holidays, while the number of men employed may have fluctuated. These fluctuations must, therefore, be included in his unit of data before the data is really comparable. This can only be done by dividing the total output by the total hours worked, thus fixing the unit as the output per man-hour. In this unit all the variable factors will be included and the manager will then be in a position to compare the output per man per hour in the respective periods, and from this comparison to make a reliable conclusion as to the relative efficiency of his factory.

Similarly, a transport undertaking cannot obtain reliable comparisons by comparing total receipts and total expenses for various periods. Fluctuating factors must be taken into account by expressing the data in units of receipts per car-mile or ton-mile and expenses per car-mile or ton-mile.

Homogeneity

The second essential characteristic is homogeneity. This means that all the facts must be analysed into groups containing only those units that are similar. Thus, if it is stated that the average earnings in the textile industry are 27s. 6d. per week, and in the building industry 48s. 4d., this information is so wide as to be practically useless. Each group contains men and women of all ages and, therefore, the data are not homogeneous. In comparing average earnings of employees either in a single factory or in particular industries, the wages of men must be considered separately from those of women, in order to secure homogeneous groups, for it is obvious that any average which includes both men and women will be affected by the proportions of men and women employed. Thus, more valuable information is given when it is shown that the average earnings of men in the textile industry are 42s. 11d. per week as compared with the average earnings of men in the building trade of 48s. 4d. per week, while the average earnings of women in the textile industry are 23s. 5d. per week, whereas no women at all are employed in the building trade. Even when the figures are presented in this form, however, the groups are not absolutely homogeneous, since each group contains workers of varying ages who are engaged in different branches of the trade, so that for a comprehensive survey, much more detailed classification is required.

A further example may be considered in which a comparison is made between the cost of production of similar articles in two factories *A* and *B*. Table I shows the total output and the cost for the two factories, and suggests that the average cost is the same for both.

TABLE I

	A	B
Total output	2,000	3,000
Total cost	£4,000	£6,000
Average cost per article . . .	£2	£2

It is found, however, that both factories are producing two classes of articles which may be called *X* and *Y*, and that, therefore, the data in the above table are not homogeneous. When the facts are classified into homogeneous groups, relating to each class of product respectively as in the second table, it is found that

TABLE II

	Factory A Article		Factory B Article	
	X	Y	X	Y
Total output . . .	1,000	1,000	1,000	2,000
Total cost . . .	£1,000	£3,000	£1,000	£5,000
Average cost per article .	£1	£3	£1	£2 10s.

factory *B* is producing one class at the same cost as factory *A*, but is producing the second article at a lower cost.

This example shows clearly how erroneous conclusions may be made when the collection of data is not divided into homogeneous groups.

Completeness

Before comparisons can be made, all the data having a bearing on the problem involved must be considered. Thus, it is not sufficient to judge the success of a business solely on the basis of the profits which have been made. Before such conclusions are made, data relating to capital involved, turnover, and similar factors must be taken into account.

When deciding upon the site for a new factory, different suggestions will be considered. It is not sufficient to make a decision solely on the cost of the purchase of the land. Consideration must be given to the rates payable in each case, and even in regard to rates the data are not complete unless the basis of assessment in each area is also known, for, although the actual rate may be lower in one district than in another, the method of assessment may be so different that the actual rates payable may be higher in the lower rated district. The cost of obtaining the necessary skilled labour in one area must be compared with that in the other, since, for example, recognized rates of wages for most trades are higher in London than in the country.

Thus, it is only after complete data have been considered that a wise decision can be taken.

Stability

Data must be stable. It is essential, when data extending over a period are under review, that such should relate to a reasonable length of time. If it is desired to purchase a business, it is not sufficient to rely on the results for one year, since this may have been exceptionally good as a result of peculiar circumstances. The results for several years must be considered in order to modify the effect of exceptional years and to reveal the true position. At the same time, the period reviewed should not be so long as to render the data useless by reason of the lapse of time. Thus, to include in the data mentioned above the earning capacity of the business twenty years ago would be absolutely valueless.

Representativeness

In many cases it is impossible to examine every example which comes under the scope of a particular inquiry, and the method of sampling may be adopted. Care must be taken, however, when this method is employed, that a sufficiently large number of samples is taken, in relation to the possible number of individual cases, to obtain returns which will vary as little as possible from the returns which would be obtained from a consideration of the whole.

Personal Bias

Little need be said to emphasize the importance of data being absolutely free from any form of personal bias. Only too frequently investigators begin their inquiries with a preconceived idea of what result must be obtained, and they are, therefore, led to overlook imperfections in their data which may fail to be comparable, homogeneous, or relative to a sufficiently long period. Similarly, two men may arrive at the value of a factory building. The one being a prospective purchaser will value every advantage the construction possesses on a conservative basis and attach great importance to its imperfections. On the other hand, the vendor will value the property on a more liberal scale, and will, as far as possible, ignore any imperfections there may be. In other words, both investigators have allowed personal bias to affect their conclusions.

Not only are the attributes mentioned essential before data can be used satisfactorily for a particular problem, but care must be exercised in drawing conclusions from any published statistics which do not always completely satisfy these requirements. Before any deduction is made from such sources, the extent to which they may be relied upon must be ascertained and reservations made accordingly.

In concluding the examination of statistical data and its compilation, we must observe one other fact, namely, that in any commercial inquiry or investigation of other kinds the problem envisaged at the commencement should be of real value. Problems should be important ones that will justify the work involved in their solution. They should not deal with subjects too insignificant to have any vital effect on the prosperity or failure of the business.

EXERCISES

1. How may the accuracy of sampling be verified?

2. Draw up a skeleton questionnaire to ascertain from 10,000 working-class families the number of earners, male and female, and the number of children and adults dependent thereon.

3. What methods of compiling the necessary data would be employed to investigate—

(*a*) The allocation of expenditure in working-class budgets?

(*b*) The extent to which acetylene welding is carried out by various concerns?

(*c*) The effect of working overtime in an individual factory as compared with the effect of working on shifts?

CHAPTER III

ACCURACY AND APPROXIMATION

THE accuracy of statistical data must now be considered, and, in this respect, the difference between statistics and accounts is important. When the accounts for the year are produced by the accountants of a company they must, of course, be absolutely accurate, but when the chairman of the company, in considering the accounts, begins to make comparisons with the accounts for previous periods, statistical methods are employed and the same degree of accuracy is not called for. Most certainly he will disregard all shillings and pence, and it is quite possible that he will refer to any of the figures only in so far as thousands of pounds are involved. It will be seen, therefore, that since statistics is largely a science of estimates, the accuracy called for in accounting is not essential to the solution of a problem. In many cases it is impossible to obtain absolute accuracy. For example, in investigations in which the method of sampling is involved, there is inevitably a discrepancy between the results obtained and those which would have been obtained had it been possible to investigate every detail.

It is, therefore, important to decide at the outset of any inquiry the degree of accuracy which is essential to give the results their maximum value. If, for example, an investigation deals with a comparison of the price of pig iron over a period of years, it is sufficient to express the data in tons, but if the price of gold is under consideration, then the unit must be the price per ounce, owing to the high intrinsic value of gold. Whatever degree of accuracy is decided upon, it is essential that this fact should be clearly stated when the data are presented in order that if the data should be used for further calculations, the accuracy of any final result will be known. Thus, the estimated population of a country may be expressed as 54,700,000 ± 100,000 (that is, between 54,600,000 and 54,800,000), 54,700,000 to the nearest 100,000 or as 55,000,000 to the nearest million. The precise degree of accuracy will depend upon the purpose for which the data will be used, and will also be controlled to a large extent by the possibility of obtaining accurate original data. For example, in the case of a small island it is possible to state the population to the nearest 10 or 100, but in the case of a large country the number of inhabitants to the nearest 500,000 may be the nearest approach to accuracy obtainable.

Approximation

It follows, therefore, that in many investigations the facts can be presented advantageously by using approximate figures, the approximations enabling the particulars to be grasped more easily, whilst at the same time, conclusions drawn from the data are precisely the same as they would have been if based on a consideration of the absolutely accurate figures.

The following table shows the actual profits and turnover of a company together with the percentage of profits on turnover.

TABLE III

Year	Profits £	Turnover £	Percentage of Profits on Turnover
1	45,631	200,432	22·8
2	73,779	365,532	20·2
3	53,711	354,632	15·1
4	65,501	397,456	16·5
5	80,902	400,637	20·2
6	90,409	599,732	15·1

It is possible, however, to approximate the figures of profit and turnover shown above, to the nearest thousand pounds, and still obtain the same results for the investigation. This has been done in the next table.

TABLE IV

Year	Profits to nearest £1,000	Turnover to nearest £1,000	Percentage of Profits on Turnover
1	46	200	23·0
2	74	366	20·2
3	54	355	15·2
4	66	397	16·6
5	81	401	20·2
6	90	600	15·0

The percentages of profits on turnover as shown by the approximated figures are so close to those obtained by using the actual figures that nothing is detracted from the value of the comparison. At the same time, the almost continuous fall in this percentage and the low figures for the third, fourth, and sixth years are equally apparent. Also, the approximated figures show more easily the actual trend of profits and turnover to increase in actual value, whilst the calculation of the percentage is facilitated by discarding figures which do not affect the final conclusions.

Approximation, therefore, enables the facts shown by data involving large figures to be shown more clearly, and facilitates further calculations, while detracting nothing from the value of the data. In other investigations the inclusion of detailed figures is misleading, suggesting a degree of accuracy that it is quite impossible to obtain. This is illustrated by the following figures relating to the Balance of Trade. To produce these figures, actual data are available concerning imports and exports compiled by customs authorities, who receive full details of all goods coming into the country, and of all goods shipped from its ports. These figures, however, constitute only a portion of the factors which go to make up the final Balance of Trade.

Other data must be included to cover the "invisible" items, but in these cases no concrete figures are available. At best they can only be estimated, and it is impossible to achieve any great degree of accuracy.

	Million £'s
Excess of imports over exports	411
Estimated excess of Government receipts from overseas	16
Estimated net shipping income	80
Estimated net income from overseas investments .	165
Estimated net receipts from short interest and commission	30
Estimated net receipts from other sources . .	10
	301
Estimated debit balance	110

In fact, the actual error which is likely to occur in these estimates is a frequent subject of controversy. The whole of the data is, therefore, approximated to millions of pounds.

The figures shown on p. 29 are largely estimates based on such facts as are available, consideration being given to the economic conditions to which they are subject. Nothing of the value of the figures is lost as a result of the approximations in which they are expressed, whilst an impossible accuracy is in no way implied.

Methods of Approximation

When data are compiled and it is decided to present the figures in approximated form, it is necessary to adopt one of several methods of approximation. Probably the most commonly used method is to approximate to the nearest whole number, or to the nearest specified decimal place. The extent to which approximation is carried will depend upon the degree of accuracy which it is desirable to retain in the data. Thus, if an inquiry relates to the production of pig iron in various areas, it will probably be sufficient to approximate the production figures to the nearest thousand tons. In this case 459,400 tons would be expressed as 459,000 tons to the nearest thousand, or 740,800 tons as 741,000 tons to the nearest thousand. In other cases even this degree of accuracy may not be called for, and it may then be sufficient to approximate these figures as 500,000, and 700,000 to the nearest hundred thousand tons respectively. Similarly, the ratio between expenses and turnover may be expressed as a percentage. If this ratio accurately calculated is

55·43278 per cent it would be necessary to retain all these figures only in rare instances. The fact that they are retained would for most purposes be confusing, particularly if it is necessary to refer to a large number of similar figures. It will, therefore, usually be found advantageous to approximate and express a figure such as the above as 55 per cent or, if slightly greater accuracy is desirable, 55·43 per cent.

A second method of approximation is accomplished by adding figures to bring the data up to the next highest whole number. When this method is adopted, the figures of pig iron production quoted above would be shown as 460,000 tons and 741,000 tons respectively, or alternatively as 500,000 tons and 800,000 tons respectively, while the percentage of 55·43278 would be presented as 56 per cent or 55·44 per cent, according to the degree of accuracy decided upon. Finally, the figures may be approximated to the next lowest whole number, and in this instance the production figures would be expressed as 459,000 tons and 740,000 tons,

TABLE V

Town	(1) Actual Population	(2) Approximated to nearest Thousand	(3) Approximated by Deleting Last 3 Digits	(4) Approximated by Raising to next Highest Thousand
1	43,432	43,000	43,000	44,000
2	21,655	22,000	21,000	22,000
3	77,431	77,000	77,000	78,000
4	78,500	78,000	78,000	79,000
5	90,501	91,000	90,000	91,000
6	40,495	40,000	40,000	41,000
7	32,671	33,000	32,000	33,000
Total	384,685	384,000	381,000	388,000

or otherwise 400,000 tons and 700,000 tons respectively. The ratio between expenses and turnover would become 55 per cent or 55·43 per cent. Table V illustrates the three methods of approximation. The actual populations of seven towns are shown in Column 1. Column 2 shows the effect of approximating to the nearest thousand; Column 3, the effect of reducing to the next lowest thousand; and Column 4, the effect of adding figures to raise to the next highest thousand.

Error in Approximation

A study of the above table will show that when the data are approximated, the population for each town and the total population for the area covered differ from the exact figures by a certain margin. This margin is defined as error. It should be noted that the use of the word "error" in this connection is to denote the difference between an approximated figure and the actual figure, and does not refer to mistakes or inaccuracies in the compilation of the data. If a comparison is made between the total population for the seven towns shown in the above table as the actual total, and the totals obtained by each method of approximation, it will be observed that in each case the error which occurs has a different magnitude. Column 2 illustrates an unbiased error, while each of the other columns represent the type of error known as a biased error.

Unbiased Errors

Errors are classed as unbiased errors when the errors resulting from individual items being approximated to the nearest whole number or nearest thousand tend to cancel one another. Thus, in Column 2 of the table above, the approximated figures for towns 2, 5, and 7 are above the actual figures, while the other items are below the actual figures. In such a case the tendency of one set of items to cause the total for the group to be higher than the exact total is offset by an opposite tendency of the other approximations. The error in the final total may, therefore, be either above or below the actual total. Owing to the tendency for errors in the individual items to cancel one another, the final error is likely to be negligible, and the greater the number of items under review, the less the relative error is likely to be, although the absolute error may increase with the number of items.

Biased Errors

When the methods of approximation employed in Columns 3 and 4 are used, the resulting error is biased. In each case the errors are all in a similar direction. In Column 4 every approximation is greater than the actual figure, and as a result of this process the errors instead of being compensatory are actually cumulative. From this it follows that the greater the number of items under consideration the greater will be the relative importance of the final error. Even with a small number of items it is so considerable that it cannot be ignored.

Measurement of Error

The error may be measured in two ways. The actual difference between the approximated figure and the exact one is referred to as the absolute error, and this may be alternatively expressed as a ratio of the approximated total, when it is then known as the relative error. These methods of expressing the magnitude of errors are made clear by Table VI, which shows the errors involved in Table V, expressed as absolute errors and as relative errors.

TABLE VI

	Col. 1	Col. 2	Col. 3	Col. 4
Total .	384,685	384,000	381,000	388,000
Absolute error		685	3,685	3,315
Relative error		$\frac{685}{384000} = \cdot 00178$	$\frac{3685}{381000} = \cdot 00967$	$\frac{3315}{388000} = \cdot 00854$

Estimating Errors

Before employing for further calculations any data in which approximations have been used, it is advisable to ascertain the amount of error involved, since this will control the accuracy of the result to be expected from use of the figures as the basis of additional investigations. In many cases the exact figures are not available and, therefore, it is not possible to ascertain the actual error, so that it is then necessary to estimate the extent to which error is possible. As already indicated, the magnitude of this error will depend largely upon the method of approximation adopted in compiling the data, so that it must first be ascertained whether the error is unbiased or biased, which, of course, will be indicated by the method of approximation shown to have been used.

In order to illustrate the method of estimating an unbiased error, assume that the data shown in Column 2 only of the above table are given. Since the data are approximated to the nearest thousand, it is known that the total will be subject to an unbiased error. Each individual item may have an absolute error of any magnitude, varying from 0 to 499, resulting from approximating to the nearest thousand, but the most likely error for any item will be the average of the numbers 0 to 499, which is 249·5, or 250 approximately. It has been found that the best estimate of the absolute error in the total when subject to an unbiased error is obtained by multiplying the average absolute error of each item by the square root of the number of items. Applying this principle to the example under consideration it has been shown that the most likely average error is 250 and to obtain the estimated error in the total this figure must be multiplied by the square root of the number of items, viz. $250 \times \sqrt{7} = 250 \times 2{\cdot}647 = 662$.

This, of course, is the estimate of the absolute error, which may be expressed as a relative error by obtaining the ratio between the estimated absolute error and the approximated total; thus the estimated relative error above would be

$$\frac{662}{384{,}000} = {\cdot}00172 \text{ or } {\cdot}172 \text{ per cent.}$$

It will be seen that this estimate compares favourably with the actual error which is known to be an absolute error of 685 or a relative error of ·00178 or ·178 per cent.

When the figures shown in Column 4 alone are considered, it is known that the approximate total of 388,000 is subject to a biased error. Each of these items has been taken to the next highest thousand, so that in each case the error created may be of any magnitude between 0 and 999, since, for example, Item 1 may have actually been 43,001 or 43,999. Again, the most likely error to occur is the average of these numbers 0 to 999, which when calculated is found to be 499·5, or approximately 500. It has already been shown that the greater the number of items under review when this method of approximation is employed, the greater the total error becomes. It is, therefore, reasonable to suppose that the best estimate of a biased error will be obtained by multiplying the average error by the number of items. The estimated error in the figures in Column 4 would, therefore, be 500 × 7 = 3,500, which proves to be approximately correct, since the actual absolute error is known to be 3,315. The relative error may be estimated as in the case of unbiased errors, by expressing the estimated absolute error as a ratio of the total of the approximated figures, thus—

$$\frac{3500}{388{,}000} = \cdot 00902 \text{ or } \cdot 902 \text{ per cent.}$$

It will be seen from the foregoing discussion of approximation that many advantages can be obtained by presenting data in approximated form. At the same time, however, due consideration must be given to the method of approximation to be employed having regard to the purpose for which the data are compiled. It must be ensured that the value of the data is not prejudiced if required for additional purposes involving further calculations. In presenting the data, the method of approximation adopted should be clearly indicated in order that the possibility of error may be taken into account before using the figures for any further investigations.

EXERCISES

1. Estimate the possible error expressed absolutely and relatively when given that the total profits of 100 companies amounts to £1,955,000 when the figures have been approximated (*a*) to the nearest thousand, (*b*) to the nearest thousand above the actual figures.

2. The following table gives the number of workers in a county over a period of six years. Columns 2–5 suggest various methods of approximating the actual figures shown in column 1. Complete the table and comment upon the relative merits of the headings used.

EMPLOYEES

Year	Actual	To nearest 1,000	000's Omitted	Round Numbers
1	4,251,639			
2	7,452,431			
3	10,542,621			
4	15,453,200			
5	14,400,500			
6	18,427,520			

CHAPTER IV

THE ARITHMETIC AVERAGE

IN compiling data for the purpose of making comparisons it is, first of all, necessary to give due attention to the particular requirements of the data, and when approximations are necessary to decide upon the method to be used. It is then necessary to devise means of expressing the relation existing between the two or more sets of facts. In many cases it is necessary to compare groups relating to a large number of items of a similar nature. For example, the investigation may relate to the earnings of men in two separate factories, each engaged in the production of similar goods. The first step in such an inquiry would be to obtain data of the earnings of the employees of each factory. If it was found that one factory employed 2,000 men and the other 6,000 men, it would obviously be impossible to compare each individual item, so that before any comparison can usefully be made it is essential to obtain a simple numerical statement which will adequately represent the whole of the data contained in each group. Even if it is desired to obtain an idea of the earnings of the employees of one factory only, it is extremely difficult to do this by a study of the pay roll showing individual earnings which fluctuate between wide limits.

In order to obtain such a numerical expression as will serve for such purposes as have been mentioned, it is necessary to find an average or type which can be relied upon to present a reasonable summary of the whole of the facts. This is most commonly achieved by using the Arithmetic Average, of which there are two classes, the Simple Arithmetic Average and the Weighted Average.

The Simple Arithmetic Average

The simple Arithmetic Average has the advantage of being extensively understood. It is obtained by dividing the total of the items under review by the number of items contained in the group. Thus, in the first factory mentioned above, if the wages bill for any given week amounted to £5,000, representing the total amount earned by 2,000 employees, the average earnings would be £2 10s. per week.

The figure gives an average weekly wage which is representative of the whole of the workers under review, and if a similar average is obtained for the second factory a comparison can be made between the average earnings of the men employed in each organization. Thus, if the average for the second factory is found to be £3 per week, it is reasonable to conclude that the men employed at this factory are better paid or more efficient than those in the first factory.

The above example illustrates the advantages and disadvantages arising from the use of the simple arithmetic average. It will be observed that it is easily calculated, and this, together with the fact that it is generally understood, is particularly advantageous when it is desired to present statistical data to individuals who are not conversant with higher mathematics. On the other hand, it cannot be said that the arithmetic average represents the earnings of any typical worker, since it does not follow that any one worker will be in receipt of the wage represented by the average. The Arithmetic Average takes into account every class of worker included in the total employees concerned; thus, the poorest type of worker, whose ability to earn is low, influences the average equally with the fast worker who is able to obtain a high wage as a result of greater efficiency.

Since each individual item is of equal importance, it is possible for extremes in either direction to offset each other.

This, however, is not entirely advantageous in some instances, since a small number of extreme items may affect the average to such an extent as to cause it to be misleading. In the illustration under consideration, the comparison leads to the conclusion that the workers of the second factory are paid a higher wage than those of the first factory. If, however, the second factory by reason of differences in organization employed two or three hundred highly skilled workmen, each in receipt of wages very much higher than those of other employees, this condition may affect the average of the whole group to such an extent as to obscure the fact that the other workmen were in receipt of wages of a lower level than those employed in the first factory. It follows, therefore, that before conclusions are made from such an average, reference should be made to the items comprising the average in order to ensure that the effect of extreme items on the average is fully appreciated.

The significance of this fact will perhaps be more clearly understood if the question of the earning power of a company is considered. When additional capital is required for a concern and an appeal is made to the public for funds, obviously a satisfactory record of profits must be presented before it is reasonable to hope that the capital flotation will be successful. If the published information asserts that with a capital of £100,000 the company has made an average profit over the past five years of £15,000, representing a return on capital of 15 per cent, the proposition would appear to be a good one. If over the five years to which the average refers the profits have been as shown below, then this conclusion is correct, since the actual profits have tended to increase from year to year.

TABLE VII

PROFITS OF THE X CO., LTD.

	£
Year 1	5,000
2	10,000
3	15,000
4	20,000
5	25,000
	£75,000

Average for five years $\frac{75{,}000}{5}$ = £15,000.

This, on the other hand, may not actually be the position.

The next table shows an alternative series of profits which may possibly represent the earning capacity of the concern, and it will be seen that the trend of events suggests a considerable falling off in profits, although the average is exactly the same as that shown by the above table.

TABLE VIII

PROFITS OF X CO., LTD.

	£
Year 1	20,000
2	20,000
3	15,000
4	15,000
5	5,000
	£75,000

Average for five years $\frac{75,000}{5}$ = £15,000.

This example shows clearly how the use of the arithmetic average may result in erroneous conclusions. No investor having data such as shown in the second table above would care to invest his capital in the concern since the profit-earning capacity shows a marked tendency to decline, but if the average for the five years only had been given, he may have been inclined to risk some of his capital in the undertaking. The average may, however, be used in many ways with great advantage. A further example exemplifying its usefulness will now be considered. If the cost of the manufacture of an article produced by a certain factory is under consideration, it may be found that the cost of production during year 5 worked out at £5 per article, whereas the cost of production during the previous year amounted to £7 per article. From this it would be assumed that the management had effected considerable savings during year 5. Other considerations, however, such as the imposition of import duties on raw material, may have rendered year 4, a year in which costs of production were exceptionally high. A more correct comparison would, therefore, be obtained by comparing cost of production in year 5 with the average cost of production over a number of previous years. The result of such a procedure would probably be as follows—

Cost of production—	£	s.	d.	
Year 1	2	10	–	per article
2	3	–	–	
3	3	10	–	
4	7	–	–	
	16	–	–	

Average cost of production . $\frac{16\ -\ -}{4}$ = £4 per article

From this it will be seen that the cost of production in year 5, although representing an improvement on the position during the previous year, was still 25 per cent higher than the average cost of production over the previous period of four years. By employing the average in this method, a far more valuable result is obtained from the investigation, since it is immediately obvious that further efforts to reduce the cost have to be made before the position can be regarded as satisfactory.

This example illustrates the use of a progressive arithmetic average, since each successive year is included and a new average calculated. It is obvious, however, that as the number of years increases, so the conditions in the earlier years will become less comparable with the current year. For this reason, the earlier years should be excluded from the average so soon as they cease to have a bearing on present statistics and a moving average substituted in order that comparisons will not be affected by the conditions existing in years which are so remote from the period under consideration as to render the data valueless.

The arithmetic average is also used extensively in vital statistics such as those relating the births and deaths in individual towns and countries. Thus, in a country having a population of 45,000,000, if it is stated that the number of births registered during the year is 99,000, this information is insufficient to enable comparisons to be made between the data for this year and that relating to previous years. It is again essential to obtain a type which makes comparisons possible, and for this purpose the arithmetic average birth rate per 1,000 of the population is calculated, and in this case would be expressed as 2·2 births per 1,000 of the population per annum. This figure can readily be compared with the average births per 1,000 of the population of other countries or of different periods. Similarly, local authorities are required to compile data of the mortality among their populations, and this data must be presented as an average number of deaths per 1,000 of the citizens before any value can be obtained from the statistics. When such data are presented, the course of the death rate for various years will reflect the incidence of various diseases, freedom from mortality in other classes of illness, the efficiency of sanitation, health, and other services.

The Weighted Average

Suppose now that the data under consideration refer to the income of the inhabitants of a town, and that the following information is available as to the respective incomes—

	£
1. One inhabitant has an income of . . .	10,000
2. One inhabitant has an income of . . .	1,000
3. Ten inhabitants have an income of . .	500
4. Three inhabitants have an income of . .	350
5. Twenty inhabitants have an income of . .	200
6. One hundred inhabitants have an income of .	100

If it is desired to ascertain the average income it may be assumed that this average could be obtained by totalling the various incomes and dividing by the number of items, thus—

$$\frac{£10{,}000 + 1{,}000 + 500 + 350 + 200 + 100}{6} = £2{,}025.$$

In order to test the accuracy of this figure, the average may be multiplied by the number of inhabitants, and if it is a correct arithmetic average, the figure obtained will represent the total income of the town. This calculation would be carried out as follows—

$$£2{,}025 \times 135 = £273{,}375.$$

The actual total income, however, is—

$$(1 \times 10{,}000) + (1 \times 1{,}000) + (10 \times 500) + (3 \times 350) + (20 \times 200) + (100 \times 100) = £31{,}050.$$

This proves that the method employed to obtain the average is incorrect and illustrates the fact that when the number of items in each group varies, then it is essential to employ the weighted arithmetic average. This average is calculated by multiplying each component item by a certain number representing the varying importance of each item, known as weights, and dividing the total of the weighted items by the sum of the weights involved.

In the above example the weighted average would be obtained by using as weights the number of inhabitants in each group—

$$\frac{(1 \times 10{,}000) + (1 \times 1{,}000) + (10 \times 500) + (3 \times 350) + (20 \times 200) + (100 \times 100)}{1 + 1 + 10 + 3 + 20 + 100}$$

$$= \frac{31{,}050}{135} = £230.$$

It will be apparent that this is the true average, since, when the figure of £230 is multiplied by the total number of inhabitants, the product is equal to the total income.

The following example will illustrate clearly the importance of using a weighted average and not a simple arithmetic average.

If the prices at which a commodity (x) is sold by a merchant are 1s., 1s. 6d., 1s. 9d., 2s., and 2s. 6d., the simple arithmetic average price would be—

$$\frac{\text{1s.} + \text{1s. 6d.} + \text{1s. 9d.} + \text{2s.} + \text{2s. 6d.}}{5}$$

$$= \frac{\text{8s. 9d.}}{5} = \text{1s. 9d. per lb.}$$

In the above calculation it is assumed that every class of x has an equal importance in his sales, whereas the actual quantities sold would vary considerably. In ascertaining the average price, this fact must be taken into account and can only be done by employing the weighted average. Having, therefore, found his sales to be 2,567 lb., 1,967 lb., 1,546 lb., 1,232 lb., and 879 lb. respectively, he must use these quantities as weights to obtain the weighted average.

$$\frac{(2{,}567 \times 1/\text{-}) + (1{,}967 \times 1/6) + (1{,}546 \times 1/9) + (1{,}232 \times 2/\text{-}) + (879 \times 2/6)}{2{,}567 + 1{,}967 + 1{,}546 + 1{,}232 + 879}$$

$$= \frac{\text{£}644\ 4\text{s.}\ 6\text{d.}}{8{,}191} = 1\text{s.}\ 6{\cdot}88\text{d. per lb.}$$

If he decides to blend a quantity of each class of *x*, he may calculate the average price, and, finding this to be 1s. 9d., decide that this must be his selling price. Should, however, his mixture be in the proportions of 10, 4, 2, 1, 1, the weighted average would be valuable to him, since it would indicate that he could afford to sell his blended commodity at a much lower figure—

$$\frac{(10 \times 1/\text{-}) + (4 \times 1/6) + (2 \times 1/9) + (1 \times 2/\text{-}) + (1 \times 2/6)}{18}$$

$$= \frac{24\text{s. } 0\text{d.}}{18} = 1\text{s. } 4\text{d. per lb.}$$

This principle is frequently employed to carry out a profitable transaction. For example, a trader may be commissioned to purchase 2,000 articles for a customer. He may find it possible to obtain 1,500 articles at the price of £1, but it may be impossible to obtain the balance of 500 at a price less than £1 10s. If, therefore, he purchases the whole quantity in this method, his average price, calculated as a simple arithmetic average, is £1 5s. The correct average is, however, the weighted average.

$$\frac{(1{,}500 \times \pounds 1) + (500 \times \pounds 1\ 10\text{s.})}{2{,}000} = \frac{2{,}250}{2{,}000} = \pounds 1\ 2\text{s. }6\text{d.}$$

This cost, calculated over the whole of his purchases, may enable him to carry out the transaction at a satisfactory margin of profit.

The Importance of Weights

In the calculation of the weighted average it is found that the result is not seriously affected by the exactitude of figures used for weights, as the error arising from the use of approximate figures is usually so small as to be negligible. This may be illustrated by calculating the weighted average price for *x* as quoted above, but employing approximated figures in place of the exact figures of sales as obtained. The calculation would then become—

$$\frac{\begin{gathered}(26 \times 1/\text{-}) + (20 \times 1/6) + (15 \times 1/9) + (12 \times 2\text{s.}) \\ + (9 \times 2/6)\end{gathered}}{26 + 20 + 15 + 12 + 9}$$

$$= \frac{\pounds 6\ 8\text{s. }9\text{d.}}{82} = 1\text{s. }6{\cdot}84\text{d. per lb.}$$

In this case there is very little difference in the average price when the weights are approximated to the nearest hundred.

While the accuracy of the weights employed does not have an important effect upon the average, any variation in the value of the items themselves is a much greater consideration. A comparison showing this effect is obtained by approximating the prices to the nearest shilling in the above calculation, but retaining the exact weights originally employed.

$$\frac{(2{,}567 \times 1/\text{-}) + (1{,}967 \times 1/\text{-}) + (1{,}546 \times 2\text{s.}) + (1{,}232 \times 2/\text{-}) + (879 \times 2/\text{-})}{2{,}567 + 1{,}967 + 1{,}546 + 1{,}232 + 879}$$

$$= \frac{£592\ 8\text{s. } 0\text{d.}}{8{,}191} = 1\text{s. } 5{\cdot}35\text{d. per lb.}$$

It follows, therefore, that in calculating a weighted arithmetic average, every consideration must be given to ensure the greatest possible accuracy of the magnitude of each item, whereas the weights may be approximated, thus considerably facilitating the ease with which calculation can be made.

This factor has also a further important aspect in that in many investigations it is not possible to obtain accurate weights. Approximate or even estimated weights may be employed without reducing the value of the investigation by reason of this fact.

An investigation into the effect of changes in the prices of food is a typical example of such a case. It is quite possible to obtain comparative data of the actual prices. For example, the prices of butter, cheese, milk, and eggs in 1913 and 1923 were as shown in the following table.

TABLE IX

COMPARATIVE PRICES OF BUTTER, CHEESE, MILK, AND EGGS

Commodity	1913 Price	1923 Price	Increase in Prices	Percentage Increases
Butter . .	1s. 2d. per lb.	1s. 8d. per lb.	6	$42\frac{6}{7}$
Cheese . .	7d. per lb.	1s. 1d. per lb.	6	$85\frac{5}{7}$
Milk . .	4d. quart	7d. quart	3	75
Eggs . .	$1\frac{3}{4}$d. each.	$2\frac{1}{2}$d. each	$\frac{3}{4}$	$42\frac{6}{7}$
				$246\frac{3}{7}$

Since the above table also shows the percentage increase for each item, the simple arithmetic average percentage increase is seen to be $\frac{246\frac{3}{7}}{4} = 61\frac{17}{28}$ per cent. This figure fails to take into account the fact that butter, cheese, milk, and eggs are not all equally important to the consumer, since, in calculating the simple arithmetic average, each item is equally important. This is obviously a case in which the weighted arithmetic average should be used, but the difficulty is to obtain the appropriate weights since every consumer will purchase varying amounts of the respective commodities. It is, however, known that in both years, the average expenditure on these goods for a normal family was approximately in the ratio of 4 : 1 : 2 : 3. These figures may, therefore, be used as weights with the following result—

$$\frac{(4 \times 42\frac{6}{7}) + (1 \times 85\frac{5}{7}) + (2 \times 75) + (3 \times 42\frac{6}{7})}{4 + 1 + 2 + 3}$$

$$\frac{535\frac{5}{7}}{10} = 53\frac{4}{7} \text{ per cent.}$$

This percentage gives a more correct impression of the importance of the increases in prices than that obtained by the simple arithmetic average, because it takes into account the relative importance of each item. Although this relation cannot be exactly calculated, it is known that since the use of approximate weights does not materially affect the average the figure obtained by using the approximate ratio of 4 : 1 : 2 : 3 is one which can be relied upon to fairly represent the importance of the increase in prices, e.g. the effect on cost of living.

Practical Application

The weighted average must be applied to many commercial problems which cannot be satisfactorily dealt with by means of a simple arithmetic average. A particularly interesting application refers to the cost of carrying out nickel plating. In this operation numerous articles of various sizes and types are plated simultaneously in the same plating vat. The salts contained in the vat and the anodes employed are used continuously for a continuing flow of work, with the result that the material used in plating any given article cannot be ascertained. In view of this fact, it is difficult to obtain the cost of carrying out the plating operation and so make accurate provision for this part of the production of any article. An ingenious method of obtaining the cost of plating based on the weighted average has been applied and proved accurate. The cost of plating is largely controlled by the area of the articles to be plated, since it consists of adding a superficial film over the area of the article. The method employed, therefore, is designed to arrive at an average cost of plating per square inch. To do this the actual total cost of the department in which the operation is carried out is carefully recorded over a period of one month, separate records being retained of materials used and labour charges. During the same month a record is also kept of the total quantity of articles which pass through the vats, in many cases of very varying character. While the total number of articles in each group is then known, it becomes necessary to weight each item in such a manner that its relative importance in the total cost of plating over the period is taken into account. To do this, the superficial area of each article in so far as it has been plated is calculated, and the area so obtained for each article is used for weighting purposes. By multiplying the area of each different article by the number of articles in each group, the total area plated is arrived at, and when this total is divided into the total cost of all the work carried out the weighted average cost of plating per square inch is obtained. This average cost can then be employed in calculating the cost of plating any other article of known area, and by comparing the cost as calculated at various periods the efficiency of the department at varying periods can be compared.

To further illustrate the weighted average, the following problem may be considered. In a certain factory operatives of varying ages are employed, and it is found that on similar work the average output per worker of each respective age is as shown in Table X, which also includes the number of workers of each age.

TABLE X
OUTPUT OF EMPLOYEES OF VARYING AGES

Ages	Average Output	Number of Workers
16	200	100
17	400	200
18	500	200
19	600	300
20 and over	600	700
		1,500

To obtain the average output it is necessary to employ a weighted average, thus—

$$\frac{(200 \times 100) + (400 \times 200) + (500 \times 200) + (600 \times 300) + (600 \times 700)}{100 + 200 + 200 + 300 + 700}$$

$$= \frac{800{,}000}{1{,}500} = 533.$$

It is found, however, that by making various alterations in the method of manufacture necessitating the

TABLE XI
OUTPUT OF EMPLOYEES OF VARIOUS AGES AFTER ALTERATION IN PROCEDURE

Age	Average Output	Number of Employees
16	—	—
17	440	100
18	550	200
19	660	400
20 and over	660	800
		1,500

substitution of older employees for the younger ones, it is possible to increase output for the other grades of workers by 10 per cent. The result of this alteration is shown in Table XI.

If now it is desired to know the effect of this alteration on the complete average, it is not sufficient to assume that this figure will also increase by 10 per cent, and a new weighted average must be obtained. This then becomes—

$$\frac{(440 \times 100) + (550 \times 200) + (660 \times 400) + (660 \times 800)}{100 + 200 + 400 + 800}$$

$$= \frac{946{,}000}{1500} = 631.$$

The importance of using a weighted average in such a case will be seen from the fact that, although the increase for each item is only 10 per cent, the final average has actually increased practically 20 per cent as a result of the altered number of workers in each grade. It will be observed that if the simple arithmetic average had been used, the average would have been 460 and 577 respectively, giving an entirely erroneous conception of the actual average and the importance of the change effected.

It may be concluded, therefore, that in many cases it is sufficient to employ the simple arithmetic average, but in other instances it is necessary to obtain the average by means of the weighted average method. It must be remembered, however, that the weighted average is simply a type of arithmetic average and is subject to the same advantages and disadvantages as the simple arithmetic average; in fact, if the number of weights involved is very numerous there will be a tendency for the weights to so counterbalance one another that the weighted average will be very little different from the simple average. This depends largely on the size of the items and on the size of the weights relating to each item. If large weights go with large items or *vice versa* the average will be seriously affected whenever weights are neglected, or when erroneous ones are used.

EXERCISES.

1. The total income of a village of 135 people is £31,050. What is the average income, and why may this average be misleading? If the total income tax paid by the inhabitants is £1,552 10s., what is the average tax in the £ paid per head?

2. The production of wheat in three counties is as follows—

(*a*) 4,376,000 acres 36·0 bushels per acre.
(*b*) 5,943,000 acres 34·6 bushels per acre.
(*c*) 7,932,000 acres 35·3 bushels per acre.

Find the weighted average for the whole area, and also show the effect of using approximated weights.

3. Below will be found the number of books sold at a certain shop at various prices. Find the average price of the books sold.

Price		No. of Books Sold
s.	*d.*	
	3	24
	9	25
1	3	80
1	9	170
2	3	155
2	9	141
3	3	92
3	9	110
4	9	80
5	9	30

CHAPTER V

THE MOVING AVERAGE

It has already been suggested that when comparisons are made between data relating to one year and similar facts relating to previous periods, it is not always possible to obtain satisfactory results by comparing the data with that for the period immediately preceding that under review. For example, the profits of a company may, for the present year, be £50,000, as compared with £70,000 for the previous year. This would suggest that the figure of £50,000 is most unsatisfactory and leads to the conclusion that the business has not been conducted in a satisfactory manner. If, however, the investigation is carried further, it may be found that the profits over the period of the company's existence were as shown in Table XII.

A study of this record immediately shows that the results of the fifteenth year's trading are not so unsatisfactory as it may at first seem, since the profits of £70,000 which resulted from the trading of the fourteenth year were exceptionally high. Probably, if the investigation is carried further into circumstances which cannot be recorded by statistical data, it will be found that these exceptionally high earnings were the outcome of certain market conditions which were not recurrent, or were derived from economies which could not be repeated.

On the other hand, the profits for the fifteenth year may be compared with the average profits over the entire life of the company. This average is £15,900, but the table of profits given below reveals the fact that this average includes the very small profits of the first few years of the company's existence, so that such a comparison is of little value in judging the trend of progress which is being made by the company. This arises from the fact that the average employed includes years which are so remote from the period under consideration as to be subject to entirely different conditions and, therefore, valueless for comparison purposes.

The problem that remains is, therefore, to show the trend of progress from the point of view of earning

TABLE XII

PROFITS OF X CO., LTD.

Year	Profits for Year	Five Years' Moving Average
	£	£
1	100	
2	100	
3	500	
4	800	
5	1,500	600
6	10,000	2,580
7	10,000	4,560
8	5,000	5,460
9	3,000	5,900
10	14,000	8,400
11	25,000	11,400
12	8,500	11,100
13	40,000	18,100
14	70,000	31,500
15	50,000	38,700
	£238,500	

Average Profits per annum over 15 years £15,900 capacity of the company in such a manner that comparisons are not affected by the results of early years. It is also essential to eliminate as far as possible the effect of exceptionally good and exceptionally poor years in order to show the underlying development which has taken place. The table indicates that there has been a tendency for profits to increase from year to year, but several years appear to have shown a falling off from this tendency. It is necessary, therefore, to disentangle the short-term fluctuations from the long-term trend in order to show clearly how this tendency has operated.

This is accomplished by means of the moving average, which is an application of the simple arithmetic average and is illustrated in Column 3 of the table. To obtain the moving average a period is taken which is long enough to eliminate the effect of fluctuations from year to year, but not so long as to include data subject to widely varying conditions. In the above example the period decided upon is a five-year term. The average is then calculated from the first five years. To obtain the moving average for the ensuing period the sixth year is included, and the first year omitted, and the average calculated over the remaining period of five years. Similarly, for each succeeding year the average is calculated over the same period, but omitting the earliest year and substituting the last year.

The moving average having been calculated in this manner, it is possible to ascertain the actual trend of progress. From Column 3 of the table which shows the moving average it will be seen that the progress has been definitely upwards. The effect of the exceptionally low years in the eighth, ninth, and twelfth years upon the moving average should be carefully noted. There appears to have been a falling off in these years of considerable magnitude since the profits for each year were below the moving average of the profits for the preceding five years. In the first two cases, however, the fall is not great enough to cause a reversal of the upward trend which is shown by the moving average, but in the twelfth year the fall is sufficiently great to cause a fall in the moving average for the period in which this year is first included. Similarly, the extraordinarily high profits for the fourteenth year, while causing the trend to move sharply upwards, are eliminated to such an extent that the more normal profits for the fifteenth year do not involve a change in the direction of the trend.

All these facts are most clearly shown when the moving average is plotted on a graph as illustrated in Diagram I. The graph also includes the actual annual profits from which it is possible to compare the profits for each year with the moving average, thus obtaining an indication of the magnitude of the fluctuations which have taken place.

As a further illustration of the moving average or trend the average annual export price of coal as shown in Table XIII may also be considered. It will be seen that while fluctuations take place from year to year there is also a general movement from which the annual fluctuations are merely deviations of a more or less pronounced extent. This general movement or trend must be emphasized by eliminating the short-term fluctuations as in the previous example, which is again accomplished by means of a five-year moving average as shown in Column 2.

Although abnormally low prices obtained in 1896 and abnormally high ones in 1900, the abnormalities are removed when the moving averages are considered. The moving average indicates the trend of events in such a way that while the effect of exceptional deviations from the main trend is modified, the importance of short-term fluctuations is readily shown by a comparison with the trend from which the effect of abnormalities has been eliminated. In this way it is possible to make conclusions which are not rendered incorrect by reason of obsolete data or violent fluctuations entering into the comparison.

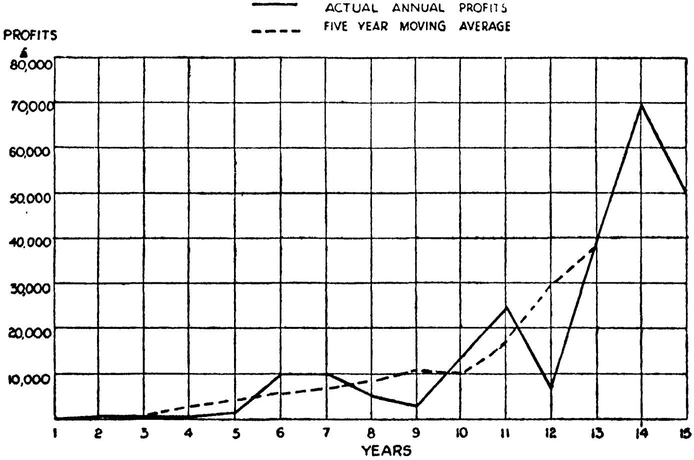


DIAGRAM I. PROFITS OF X COMPANY, LTD.

When the facts shown in Table XIII are presented graphically, the angles and slight fluctuations may be

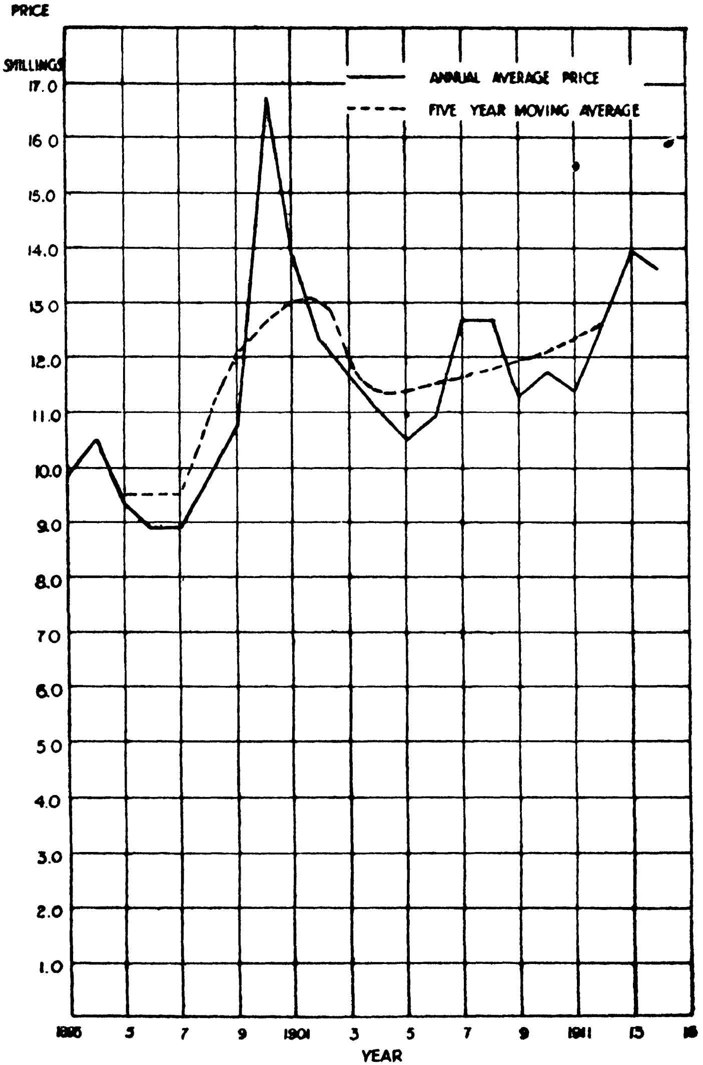


DIAGRAM 2. ANNUAL AVERAGE PRICE OF COAL

TABLE XIII

ANNUAL AVERAGE EXPORT PRICE OF COAL

Year	Column 1 Annual Average Export Price of Coal	Year	Column 2 Five-year Moving Average Export Price of Coal
1893	9·90	—	—
1894	10·50	—	—
1895	9·33	1893–1897	9·51
1896	8·85	1894–1898	9·52
1897	8·98	1895–1899	9·56
1898	9·92	1896–1900	11·04
1899	10·72	1897–1901	12·05
1900	16·75	1898–1902	12·71
1901	13·86	1899–1903	13·06
1902	12·29	1900–1904	13·15
1903	11·70	1901–1905	11·91
1904	11·13	1902–1906	11·32
1905	10·56	1903–1907	11·41
1906	10·90	1904–1908	11·62
1907	12·75	1905–1909	11·66
1908	12·77	1906–1910	11·89
1909	11·30	1907–1911	11·99
1910	11·72	1908–1912	11·98
1911	11·43	1909–1913	12·22
1912	12·70	1910–1914	12·69
1913	13·94	—	—
1914	13·65	—	—

smoothed out as shown in Diagram 2, since they are accidental and unimportant. The graph thus indicates the long-term trend of the price of coal, which it will be seen is upward until 1902, when it falls rapidly, but again begins to rise slowly after 1904. This trend enables a forecast to be made as to the possible direction of prices in the immediate future, and also shows clearly those years in which prices have been widely divergent from the normal tendency.

Commercial Application

For the business man the trend is important in its various applications, and probably the most valuable is in respect of turnover. If an undertaking is to prove prosperous it is essential that the turnover should show a tendency to increase. Individual years will fluctuate in so far as they are affected by particular influences so that it becomes imperative to ascertain the moving average of the turnover if a true conception of the progress made is to be obtained.

In addition to this, a study of the trend of turnover over past years enables the management to forecast with reasonable possibility of success the probable trend of sales in the near future. This is a valuable guide to future policy and forms the basis for budgetary control of expenses and material purchases.

Statistics are also available for the majority of industries as a whole. Thus, if a manufacturer of electrical machinery finds his own turnover is showing a tendency to decline while the industry as a whole is progressive, there is obviously an indication that something is wrong with his own organization.

Not only must the turnover of a business be studied in this manner, but the trend of expenses must also be similarly controlled. While it is vitally important that attention should be given to turnover, it is equally important to watch expenses to ensure that the upward tendency of all classes of expenditure does not exceed the progress made by sales, as in such a case the percentage of profit on turnover will be correspondingly reduced. If the expenses for any one year when considered in relation to the trend over a number of years shows an abnormal increase, investigations must be made to ascertain the factors which have caused the deviation. All items of expenditure having been investigated, steps can be taken to reduce expenses in any ensuing period.

Sufficient attention is rarely given to the application of the moving average to the price of raw materials, yet by studying prices of essential raw materials in this manner considerable savings can be effected. The moving average prices of the raw materials will enable a manufacturer to see the underlying trend of prices of any material in which he is interested. This will provide a guide to the direction in which prices are likely to move, so that with this information available it is possible to formulate a wise purchasing policy. An additional study of the short-term fluctuations which take place in the prices of any material will indicate the seasonal fluctuations in prices, so that it then becomes possible to decide upon those periods at which prices tend to fall to their lowest point indicating the most favourable seasons for purchases to be made.

Finally, it should be remembered that the trend of production of raw materials, particularly of such basic materials as iron or coal, is significant of the condition of business in general. In fact, this trend is one of the most valuable barometers of trade conditions which combined with such factors as the movement of raw material prices, the price of industrial securities and monetary conditions enable estimates to be made of business activity. In the study of such conditions, it is obvious that short-term fluctuations must be eliminated in order to arrive at a true conclusion concerning the underlying trend of events, and, for this purpose, it is essential, therefore, to employ the moving average. The value of such investigations cannot be over-estimated, since it is only by reference to prevailing conditions that a reasonable forecast of future tendencies can be made. On the accuracy of such forecasts largely depends the success of individual businesses whose policies must be organized to take advantage of any improvement that may be imminent in general trade conditions. Frequently the marketing of a new product depends largely for its success on the market conditions operative at the time the marketing is commenced. A favourable opportunity must, therefore, be chosen and the decision as to when this opportunity occurs can only be made after a careful study of the trend of prices, business activity, and general trade conditions.

EXERCISES

1. What is the value of the moving average? How may the extent of short-term fluctuations be shown?

2. Calculate five-year moving averages of the profits of a company as shown by the following table. Comment on the trend shown.

Year	Profits	Year	Profits
	£		£
1	10,000	10	30,000
2	8,000	11	20,000
3	7,000	12	25,000
4	12,000	13	40,000
5	25,000	14	50,000
6	16,000	15	30,000
7	24,000	16	35,000
8	18,000	17	10,000
9	35,000		

CHAPTER VI

THE MEDIAN AND THE MODE

WHEN the average has been obtained by any of the methods so far described, a considerable amount of information will be available concerning the groups of data under consideration. The average may provide a means whereby it is possible to compare to a certain extent the characteristics of those groups which are similar. It is possible, however, that the average alone may not be a sufficiently good summary of the whole of the facts. For example, the arithmetic average of two groups may be identical as in the case of the average earnings for the two groups of workers included in Table XIV, although a closer examination of the figures themselves reveals distinct differences in the incidence of the figures producing the respective averages. A slight examination reveals that such an average, though comparatively simple and easily understood, is not by itself the most satisfactory method of measurement in such a case.

It is, therefore, necessary to use additional means of comparison, one of which is the median.

The Median

The median may most simply be defined as the middle point in an array of figures, set out in order of magnitude so that the number of items above the median and the number below it are equal. Thus, in

TABLE XIV

Comparison of Earnings of Two Groups of Workers

Group A				Group B			
	Earnings				Earnings		
	£	s.	d.		£	s.	d.
1	3	–	–	1	5	–	–
2	4	–	–	2	1	10	–
3	2	10	–	3	1	–	–
4	3	–	–	4	5	–	–
5	3	5	–	5	1	–	–
6	2	15	–	6	2	–	–
7	3	10	–	7	2	10	–
8	2	15	–	8	8	–	–
9	2	10	–	9	2	10	–
10	3	–	–	10	2	–	–
11	3	5	–	11	1	15	–
12	3	–	–	12	2	5	–
13	2	15	–	13	3	–	–
14	2	10	–	14	1	10	–
15	2	10	–	15	1	–	–
16	3	10	–	16	7	–	–
17	3	5	–	17	4	10	–
18	3	5	–	18	4	–	–
19	3	–	–	19	1	10	–
20	3	–	–	20	1	–	–
21	2	15	–	21	5	–	–
Total earnings	63	–	–	Total earnings	63	–	–
Arithmetic average	3	–	–	Arithmetic average	3	–	–

a simple series such as 1, 1, 2, 5, 7, 9, 11, 15, 16, 17, 18, 20, 24, 29, 40, the median is the eighth item, since the whole series consists of fifteen items, so that there are seven items above and seven below the middle point.

Applying this definition to the data given in the above table, it will be seen that for each group the median is the eleventh item after the data have been assembled in order of magnitude. This may be carried out in respect of Group A as follows—

GROUP A—ARRANGED IN ORDER OF MAGNITUDE

	£	s.	d.
4 men earn	2	10	–
4 men earn	2	15	–
6 men earn	3	–	–
4 men earn	3	5	–
2 men earn	3	10	–
1 man earns . . .	4	–	–

The eleventh item is included in the group of six men, earning £3 per week, and, since the earnings of all the men in this group are similar, the median can immediately be stated to be £3.

The second group arranged in order of magnitude is shown below—

GROUP B—ARRANGED IN ORDER OF MAGNITUDE

	£	s.	d.
4 men earn	1	–	–
3 men earn	1	10	–
1 man earns . . .	1	15	–
2 men earn	2	–	–
1 man earns . . .	2	5	–
2 men earn	2	10	–
1 man earns . . .	3	–	–
1 man earns . . .	4	–	–
1 man earns . . .	4	10	–
3 men earn	5	–	–
1 man earns . . .	7	–	–
1 man earns . . .	8	–	–

In this case the eleventh item is represented by the one man earning £2 5s. per week, which is the median for the group.

A study of these two results will reveal the effect of using the median. In the first place, the median is usually, though not always, an actual example contained in the data under review, and is capable of being actually located. This, of course, is rarely the case with the arithmetic average, which may be different from any item contained in the data.

The example considered above also illustrates the most important consideration in regard to the median—that is, that it is not affected by extremes. The arithmetic average obtained for group A may be said to fairly represent the whole of the group, since the whole of the items deviate only within slight limits of the average. This is not so in the second group, since in this group there are a number of extremely low items which in the calculation of the arithmetic average are offset by the effect of a comparatively small number of extremes in the other direction. The median, however, is not affected by either extremes, and for this reason gives a fairer representation of the whole of the data.

While the median has thus considerable advantages, it suffers from the disadvantages that it cannot be calculated from any given formula, and that in order to ascertain the median the whole of the data under review must be arranged in order of magnitude, which in extensive investigations entails a large amount of work. By reason of these facts it follows that the median cannot be used for the purpose of carrying out additional calculations, since its significance arises from its central position in a series arranged in order of magnitude.

The Mode

When data has been arranged in order of magnitude, it is frequently found that a certain magnitude occurs more frequently than any other, so that a predominant group can be defined between certain limits. This value in any data which occurs most frequently is known as the mode, and it follows from this definition of the mode that if items are chosen at random the most likely item to appear is the modal value. The mode can be clearly seen in the following diagram, which

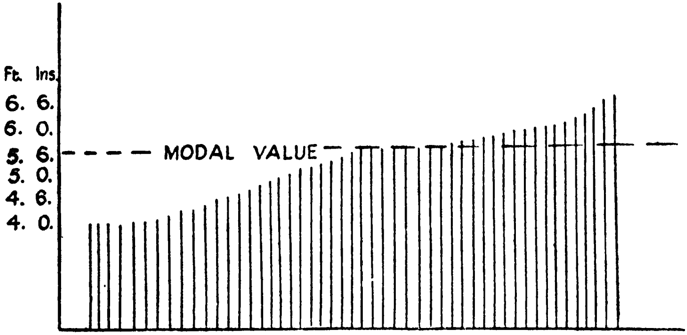


Diagram Indicating Effect of Mode

represents the heights of a group of men, each represented by a vertical line drawn to scale. The diagram includes a group of fifty items, which have been arranged in ascending order of magnitude and shows clearly the effect of the height of 5 ft. 6 in. predominating in the group. There are eight items representing this magnitude, whereas the other items range from 4 ft. 0 in. to 5 ft. 5 in., and 5 ft. 7 in. to 6 ft. 6 in. without any other magnitude predominating. The effect of this predominance of one group causes the curve obtained from plotting the various magnitudes to flatten out, thus indicating the position of the mode.

Locating the Median and the Mode

To further illustrate the median and the mode and, the method of ascertaining the magnitude of these types, consider the following table, which has been compiled from the monthly records of bankruptcies for a certain period. The data has been graded in groups according to the magnitude of the bankruptcies for each month, and the number of occurrences in each group, or frequencies, are also shown.

The data will be seen to relate to a total of fifty items, and since the median must be the case above and below which there are an equal number of items, this will lie between the twenty-fifth and twenty-sixth items.

TABLE XV

TABLE OF BANKRUPTCIES

No. of Group	Group Representing Number of Bankruptcies per Month	Number of Items in each Group (Frequency)	No. of Group	Group Representing Number of Bankruptcies per Month	Number of Items in each Group (Frequency)
1	230–240	2	9	311–320	4
2	241–250	1	10	321–330	3
3	251–260	1	11	331–340	5
4	261–270	7	12	341–350	2
5	271–280	3	13	351–360	1
6	281–290	1	14	361–370	3
7	291–300	6	15	371–380	1
8	301–310	9	16	381–390	1
	Total number of items	30		Total number of items	20

Both the twenty-fifth and twenty-sixth items fall within the eighth group representing 301–310 bankruptcies per month and including a total of nine items.

As it is not possible from the above table to ascertain the exact magnitude of the items required, it must be assumed that the magnitudes of the nine items included in the eighth group in which they are situated are evenly distributed between the minimum magnitude of 301 bankruptcies per month and the maximum number of 310.

Assuming this to be true, the twenty-fifth item is then the fourth item in this group, and its magnitude may be estimated by adding to the maximum magnitude in the next lowest group a proportionate amount of the interval contained by the group in which the item lies.

The calculation is, therefore—

$$300 + (\tfrac{4}{9} \times 10) = 300 + 4{\cdot}4 = 304{\cdot}4.$$

Similarly, the twenty-sixth item is the fifth in the group, and applying the same principle its magnitude may be estimated as

$$300 + (\tfrac{5}{9} \times 10) = 300 + 5{\cdot}5 = 305{\cdot}5.$$

The median has already been stated to lie between the twenty-fifth and twenty-sixth items, and, since the magnitude of these has been estimated as 304·4 and 305·5 respectively, it may be stated that the median lies between 304·4 and 305·5, or approximately 305 bankruptcies per month.

This figure provides a type which summarizes the whole of the data included in the table, and may, therefore, be used for purposes of comparison with any other period. Again, it will be noted that the magnitude of the median is not affected by the extremely high number of 381–390 bankruptcies per month which have been recorded for one occasion, or any other extremes in other directions.

To obtain the mode it is essential to ascertain which group contains the greatest number of items and predominates markedly over the other groups. The table shows that the eighth group has a frequency of nine, which is considerably greater than that of any other group. It may reasonably be stated, therefore, that the modal value can be expressed as 301–310 bankruptcies per month.

Size of Group	Frequency	(*a*)	(*b*)	(*c*)	(*d*)	(*e*)
5–0	50	102		162		
5–1	52		112		182	
5–2	60	130				195
5–3	70		135	210		
5–4	65	140			200	
5–5	75		135			209
5–6	60	134		207		
5–7	74		147		221	
5–8	73	147				217
5–9	74		144	212		
5–10	70	138			190	
5–11	68		120			
6–0	52					

In this example, the predominating group is so marked that it is easily distinguished, but in some series this is not the case. It then becomes necessary to widen the groups, considering together the items in two or more groups, amalgamated together in various combinations, until magnitudes between certain limits can be distinguished as predominant. To illustrate this method of determining the mode, the first column of the preceding table may be studied. This represents the number of men of given heights in a group of 843 men. No particular height is so predominant as to stand out definitely as the mode, so that it is necessary to widen the limits of the groups by amalgamation. This is done by amalgamating every two groups as in (*a*) and (*b*), and by amalgamating every three groups as in (*c*), (*d*), and (*e*).

It will be seen that in columns (*a*) and (*b*) every possible amalgamation of two groups is made, and in columns (*c*), (*d*), and (*e*) every possible combination of three groups. If the grouping which shows the highest frequency in each case is extracted, the following summary is compiled—

In group (*a*) the maximum frequency is in group 5 ft. 8 in.–5 ft. 9 in.
,, (*b*) ,, ,, ,, ,, ,, 5 ft. 7 in.–5 ft. 8 in.
,, (*c*) ,, ,, ,, ,, ,, 5 ft. 9 in.–5 ft. 10 in.–5 ft. 11 in.
,, (*d*) ,, ,, ,, ,, ,, 5 ft. 7 in.–5 ft. 8 in.–5 ft. 9 in.
,, (*e*) ,, ,, ,, , ,, 5 ft. 8 in.–5 ft. 9 in.–5 ft. 10 in.

Here groups 5 ft. 8 in. and 5 ft. 9 in. each occurs four times, and since the frequencies are somewhat irregular, the most that can be said is that the mode appears to be approximately 5 ft. 8 in. or 5 ft. 9 in.

It will be noted that the less predominant the modal value is, the more difficult it is to ascertain its magnitude, so that, in many cases, it is possible to state the mode only between fairly wide limits.

Commercial Value of the Median and the Mode

In actual practice the median and the mode are not so extensively used as may be possible. This is probably due to the fact that they are not so generally understood as the simple average. Nevertheless, both types may be applied in commercial problems. In the case of the median it is frequently overlooked that it can be applied when the factors under consideration do not lend themselves to exact measurements. For example, efficiency rating systems are becoming increasingly significant in the organization and control of personnel. Individual members are graded in accordance with the extent to which they possess those characteristics necessary for suitable and valuable employees. Yet these attributes cannot be exactly measured. It will be seen, therefore, that the principle of the median may be advantageously applied in such investigations, since one type may be decided upon as the median and the remainder of the employees grouped round it in their relative positions.

In determining wage rates it frequently happens that too much stress is laid upon the average earnings possible to the workers, which average may be seriously affected by extremes. This may be obviated by taking into account not only the arithmetic average but also the mode in checking the accuracy of piece rates fixed by means of technical processes. In this way, since the mode is not affected by extremes and represents the greatest number of workers who have similar earning capacities, a more accurate result can be obtained.

The mode may be usefully employed in relation to the output of machines. Once the modal output per machine has been ascertained, it is reasonable to expect that similar machines should be capable of producing a similar output. In this way the output of every machine can be compared with the modal output previously ascertained, and any falling off can then be investigated. Similarly, the mode will supply a valuable guide to the output most likely to result from additional machines of an identical character, and, therefore, correct estimates can be made of the additional machines necessary to produce a given increase in the capacity of a factory. The same remarks apply equally to the purchase of entirely new machinery, since the decision to purchase machines should not be based upon the speed of the machines only, but upon a fair estimate of the modal output which they are likely to give. By the application of the mode to such manufacturing problems as these, it is possible to secure the utmost efficiency from plant and machinery, with corresponding reductions in overhead charges and the curtailing of unnecessary expenditure.

The mode may also be applied to manual output in order to ascertain the conditions most conducive to the maximum efficiency of the workers, and to securing congenial and healthy working conditions.

EXERCISES

1. The following data relate to the earnings of a group of female workers. Calculate the arithmetic average and ascertain the median—

Wages	Number of Women
s. *d.*	
14 –	650
17 –	432
18 –	500
19 –	271
20 –	250
21 –	190
22 –	184
23 –	179
24 –	165
25 –	55
26 –	65
27 –	102
28 –	73

2. If the modal output for a certain machine is ascertained to be 500,000 parts per day of eight hours, comment upon the following actual output obtained for four such machines in similar working periods—

(000 omitted)

	Days						
Machine	1	2	3	4	5	6	7
No. 1	498	502	503	501	497	496	504
No. 2	450	430	400	400	375	410	430
No. 3	550	540	555	560	540	530	555
No. 4	530	430	410	550	560	500	400

CHAPTER VII

INDEX NUMBERS

THERE are many statistical problems which it is not possible to solve by means of simple comparisons between various series, and where averages or types are of no use in their solution. It is, therefore, necessary to render numerical statements comparable by reducing them to a common denominator. This can only be done by the scientific method which has been evolved whereby the fluctuations in data of various classes are ascertained by means of compiling index numbers. The method will be most clearly understood by reference to the application of index numbers to the variations in the purchasing power of money. The purchasing power of money fluctuates in accordance with the amount of money and credit available, but the fluctuations which take place cannot be measured by themselves. They are, however, reflected in the price of commodities in so far as these prices increase when the purchasing power of money decreases, and decrease as this purchasing power increases. It is not possible, however, to take one commodity and by investigating the course of prices for this commodity alone make any conclusions concerning the value of money, since each commodity is not only subject to the general influence of monetary conditions, but is affected by other considerations to a greater or less extent. For example, the price of wheat is influenced by weather conditions, prospective or actual crops, changes in customs of the consumer, while the price of such a commodity as coal is the result of a combination of factors, including demand, output restriction, and labour conditions. It follows, therefore, that it is only when widely differing commodities show similar trends in price that peculiar influences can be said to be of secondary importance and the general trend to be attributable to the fluctuations in the purchasing power of money.

To investigate the general change in prices, all special causes of variation must be eliminated as far as possible, and this can only be done by averaging together the price of as many commodities as possible from the whole of the commodities in general use. Since prices are expressed in different manners for different commodities, coal and steel as so much per ton, wheat per bushel, sugar per hundredweight, they can only be combined by converting each price into a number which is proportionate in value to a base number representing the price ruling at a given period. By averaging the resultant common denominators, an index number of the prices ruling for the commodities under consideration is obtained.

By obtaining indices at various periods the tendency for prices to rise or fall may be ascertained. If, in addition to this, index numbers are compiled for individual commodities, it is possible to decide how far the course of prices of the individual commodities or certain groups of commodities conform to the general trend; that is, to what extent they have been affected by monetary factors, and how far they have been influenced by conditions peculiar to themselves.

Compiling Index Numbers

The precise method of forming an index number may be illustrated by the following example. Data are given relating to the price of sugar, tea, and coffee over a number of years, and it is desired to compile an index number for these commodities to show the trend of prices for these articles as a whole. It is first essential to decide upon a base period, and for this purpose the first year, 1924, may be taken as the base. In some cases, it is desirable to fix the base period as the average over a number of years, but since the data refer only to nine years, it is more convenient to employ the first year's figures for this purpose.

TABLE XVI

PRICES OF SUGAR, TEA, AND COFFEE FOR THE YEARS 1924–32

Year	Sugar per lb.	Tea per lb.		Coffee per lb.	
	d.	*s.*	*d.*	*s.*	*d.*
1924	10	2	6	2	1
1925	11	2	9	2	3
1926	9½	3	–	2	4
1927	9	2	9	2	2
1928	9	2	6	1	11
1929	8½	2	6	1	10
1930	8	2	3	1	10
1931	7	2	–	1	9
1932	5	1	6	1	9

From the above it is impossible to decide upon any tendency to rise or fall over the period for the commodities in question.

In order to compile the index numbers which will enable this information to be obtained, the prices for each commodity in 1924 which has been decided upon as the base year, are expressed as 100, and the price for each succeeding year calculated proportionately to this figure. Thus, if the price of 10d. per lb. for sugar in 1924 is represented by 100, then the calculation of the proportionate figure for 1925 is as follows—

$$\frac{11 \times 100}{10} = 110.$$

The price for each year and for each commodity is similarly calculated, and by this means the prices are all converted into similar denominations. Finally, the average of the three figures obtained for each year is calculated, and an index number is then obtained for the prices of sugar, tea, and coffee for each year from 1924 to 1932 on the basis of 1924 = 100.

This is shown in Table XVII.

TABLE XVII

INDEX NUMBERS FOR TEA, SUGAR, AND COFFEE, 1924–32

1924 = 100

Year	Sugar	Tea	Coffee	Average Representing Index Numbers
1924	100	100	100	100
1925	110	110	108	109·3
1926	95	120	112	109·0
1927	90	110	104	101·3
1928	90	100	92	94·0
1929	85	100	88	91·0
1930	80	90	88	86·0
1931	70	80	84	78·0
1932	50	60	84	64·6

The indices above enable the following conclusions to be drawn. First, the combined index numbers show that the prices as a whole increased sharply in 1925, fell slightly in 1926, and from then onwards continued to decline to a very low figure in the final year. Secondly, the individual commodities themselves showed slight variations from this general tendency. For example, the prices of tea and coffee continued to rise until 1926, whereas the price of sugar began to fall in this year. Also, the rate of decline in the various commodities fluctuated slightly in the ensuing years from the combined index numbers. ”

Weighted and Unweighted Index Numbers

It will be appreciated that, in order to obtain a general index of prices which will be sufficiently comprehensive to indicate the trend of monetary values, a much larger number of articles must be included in the index. In addition, some consideration must be given to the relative importance of each item included. This particularly applies to an index designed to enable comparisons to be made between the cost of living at various periods. In such an index the component items consist of articles of general consumption, but each item has a different importance from the point of view of the household budget, and this fact must be taken into account in compiling the index. For example, rent is a far more important item than tobacco, and only when the index takes into account the relative importance of these differences can it be relied upon adequately to represent the position.

The method of weighting employed to achieve this object may for simplicity be illustrated by again referring to the index for sugar, tea, and coffee previously examined. If it is decided that sugar is three times and tea twice as important as coffee, the factors making up the index may be weighted as to 3 : 2 : 1. The proportionate numbers obtained for each item are multiplied by the weights decided upon, and for each year the total of the products is divided by the total of the weights in order to obtain a weighted average. In this manner the average is influenced to a greater extent by the figures for sugar than by taking an unweighted average, and to a slightly less degree by the figures for tea than by the figures for sugar.

The weighted index numbers as calculated in this manner are given in Table XVIII.

TABLE XVIII

WEIGHTED INDEX NUMBERS FOR SUGAR, TEA, AND COFFEE

1924 = 100

Year	Sugar	Tea	Coffee	Average or Weighted Index Numbers
1924	300	200	100	100·0
1925	330	220	108	109·7
1926	285	240	112	106·1
1927	270	220	104	99·0
1928	270	200	92	93·7
1929	255	200	88	90·5
1930	240	180	88	84·7
1931	210	160	84	75·7
1932	150	120	84	59·0

The effect of weighting the index number is most clearly seen in the index for 1932. The weighted index for this year is 59·0 as compared with an unweighted index of 64·6, this difference being due to the heavy fall in the price of sugar in the final year, which in the weighted index is given increased importance.

Advantages and Disadvantages of Index Numbers

Index numbers have the great advantage of providing a means whereby the trend of events relating to certain data can be studied, whereas this would not be possible by any other means. For example, it would be impossible to study the trend of prices of commodities, industrial and other securities, business activity, or cost of living without the aid of index numbers. They are, however, subject to criticism in various respects.

The first of these is the fact that to compile any index a base period has to be decided upon, and this may be an exceptional period. For example, 1924 has been chosen as the base in the example given above, and from the index arrived at it may be concluded that the fall in prices indicated was serious. It may be, however, that prices were abnormally high in the base year and that such a fall was inevitable. In studying data, therefore, by means of index numbers, any relevant information concerning the base period must be taken into account.

Secondly, the index number depends mainly upon the commodities which are included and the precise method of compilation adopted. It is obviously impossible to include every commodity, and the most important must be chosen, with the result that two index numbers for the same period will differ as a result of the different commodities included.

Also, it is obviously difficult to obtain standard prices for the commodities which it is decided to include. For every commodity many different prices will be in operation at one and the same time, and, in order to obtain as accurate results as possible, an average of the known prices must be taken as the ruling prices. Consideration must also be given to the variation in prices between different areas, e.g. in London and the provinces.

Finally, it is a difficult matter to decide upon the weights which are to be given to the component items. This, however, although having an effect upon the ultimate accuracy of any results obtained, does not materially affect the main conclusions, since divergence in the weights does not seriously affect the averages. It follows, therefore, that it is not possible to compile an index number for which can be claimed absolute accuracy, and that, therefore, an index must be considered only as a reasonable approximation to the conditions which it is intended to investigate.

Published Index Numbers

There are many index numbers which are regularly published, and as they are important it is necessary to survey the methods employed in compiling them.

Index of Prices. One of the best-known index numbers is Sauerbeck's *Course of Average Prices.* For this index, 45 commodities are taken, and, in some cases, the arithmetic average prices of various qualities of the same commodity are taken in order to include divergent classes of the same article. The proportionate number for each commodity is calculated by simple proportion, and the average gives the index number on the basis of 1867-1877 = 100, which is chosen as the base period.

"Statist" Index Numbers. The *Statist* Index Number is compiled in exactly the same manner as in the case of Sauerbeck's *Course of Average Prices.* In order to give additional information, the data have been classified into groups. Separate index numbers are shown for each group, so that the trend of prices of the commodities included in each group can be clearly indicated in relation to the general trend. The groups are composed as follows—

1. Foodstuffs — 19 commodities
    - (a) Vegetable foods
    - (b) Animal foods
    - (c) Sugar, coffee, and tea
2. Materials — 26 commodities
    - (a) Minerals
    - (b) Textiles
    - (c) Sundries

"Economist" Index Numbers. To obtain this index number, 44 commodities are taken, and the base period taken is the average price ruling during 1927. The selected commodities fall within the following classification: (1) Cereals and meat, (2) other foods, (3) textiles, (4) minerals, (5) miscellaneous, and separate indices are given for each group. Some commodities are given more than one quotation, so that by this means these commodities are weighted as far as possible in accordance with their relative importance. The final index is obtained by converting current prices into percentages of the base period and dividing the total for the whole by the number of commodities.

Cost of Living Index Numbers. Changing conditions and changing prices have the effect of varying the cost of living. If a standard of living is to be maintained at a certain level, the cost of maintaining that standard increases when the retail prices of commodities such as food and clothing increase, and *vice versa.* During the years since the beginning of the first Great War when prices increased considerably, it was important to obtain some measure of the increase in the cost of living and to do this index numbers were compiled, the most important being the Ministry of Labour *Cost of Living Index.*

This index takes into account the following items of expenditure: (1) Food, (2) rent, (3) clothing, (4) fuel and light, (5) sundries, and is compiled in the following manner. Careful inquiries are made each month from retailers over a wide area to obtain the average ruling prices for each commodity. This information is analysed to show the percentage increase in prices for each of the five groups mentioned above.

In the case of the first group—food—information is obtained for all important foods, for example, beef, mutton, fish, flour, bread, tea, sugar, etc., and each item is weighted in accordance with its estimated relative importance as shown by an analysis of working-class budgets made by the Board of Trade. In this way, a proportionate figure is obtained for foods, and similarly for each other item. After these investigations have been carried out, it remains to combine all the details into one comprehensive index number, which is achieved by obtaining a weighted average of the figures shown for each group. The weights used are in accordance with the estimated relative importance of each group in the average pre-war expenditure, and are: Food, 7½; Rent, 2; Clothing, 1½; Fuel, 1; Other items, ½.

The index so produced enables a measure to be made of the cost of providing the average requirements of the normal household. The figure obtained has formed the basis of fluctuations in bonuses paid to Civil Service and certain Municipal workers, and has formed the basis of adjustments in the wages of many other classes of workers.

Other Index Numbers

Frequently, index numbers are employed to express the course of events that cannot otherwise be numerically stated, valuable information being obtained in this way. Two interesting examples may be quoted—The Index of National Prosperity and also the Index of Business Activity published by the *Economist.* This latter is an attempt to combine a number of business indices into one figure showing the trend of business activity as a whole. The groups included take into account employment, power consumption, domestic trade and transport, financial factors and foreign trade, and each item is weighted according to its estimated relative importance.

It will be seen that index numbers can be applied to many commercial problems of a general nature, and may also be extended to the use of individual business problems. Wages matters may be dealt with after reference to reasonably reliable indications of the actual cost of living which determines the real wages received. A consideration of indices of prices is of assistance in formulating purchasing policy, and frequently may be applied as a guide to the conduct of sales campaigns. At times, when prices are shown to be rising, increased prosperity can usually be anticipated, and, therefore, the index numbers available enable decisions to be made as to the most opportune time to prepare for increasing turnover or to exploit a new product.

EXERCISES

1. From the data given below calculate the index number for food on 28th February, 1925, on the basis of 1914 = 100. The price of each item must be expressed as a percentage of its 1914 price, and the weight applied to these percentages—

	Weight	Average Price per lb.			
		July, 1914		Feb. 28, 1925	
Beef, *British*—		*s.*	*d.*	*s.*	*d.*
Ribs	12		10	1	6
Flank	12		6		10
Beef, *Chilled* or *Frozen*—					
Ribs	12		7		11
Flank	12		5		6
Mutton, *British*—					
Legs	6		10	1	8
Breast	6		6	1	–
Mutton, *Frozen*—					
Legs	6		7	1	–
Breast	6		4		6
Bacon	19		11	1	6
Flour	20		10	1	8
Bread	50		6		11
Tea	22	1	6	2	6
Sugar	19		2		4
Milk (per quart)	25		3½		7
Butter	20	1	2	2	2
Cheese	10		9	1	2
Margarine	10		7		7
Eggs (each)	19		1		2
Potatoes (per 7 lb.)	18		4½		10

2. Compute unweighted index numbers showing the change in market values of the following industrial shares. (Only five

items are included in this exercise, but at least twenty would require to be used to indicate general movements.)

Share	Average 1913 Prices		April, 1922		April, 1923		April, 1930	
	s.	*d.*	*s.*	*d.*	*s.*	*d.*	*s.*	*d.*
Y and Y, Ltd.	60	–	46	–	65	–	64	–
Associated Engineers, Ltd.	14	–	20	–	22	–	15	–
Coals, Ltd.	9	6	24	–	29	–	24	–
Electricity, Ltd.	16	–	23	–	20	–	22	6
Machinery, Ltd.	40	–	18	–	20	–	14	–

CHAPTER VIII

TABULATION

In the course of all investigations, the data having been collected and the method of comparison decided upon, there remains the problem of the method of presentation. Facts must be presented in such a manner that their significance and the bearing they have on the problems under consideration are shown as clearly as possible. This, in many cases, is far from being a simple matter, as it is often exceedingly difficult to decide upon the exact form to be adopted in any particular case.

Data may be presented in tabular, graphic, or diagrammatic form, and each method requires to be carefully considered in order to secure the best possible application of whichever method is adopted.

Tabulation

The function of tabulation is to arrange and amalgamate in a form which facilitates the interpretation of all the relevant information gathered together for the purpose of solving the problem under investigation. Thus, any data collected with a view to giving the position of trade in general will be expected to reveal the trend of export trade to various markets, the position of stocks held in the country, the development of new imports, the ratio of raw materials to manufactured goods included in the imports, and the volume of entrepôt trade which is transacte

On the other hand, the data compiled for individual business requirements will, of necessity, be of a much simpler nature. In this case also, however, it is frequently advisable to tabulate in subsidiary tables all the information obtained in order to compile a clear final analysis. It is also desirable to retain the detailed summaries, not only for further reference purposes in respect of the problem immediately under consideration, but also because such information is frequently valuable in connection with other problems of a similar nature.

It is obvious, therefore, that it is impossible to lay down concrete rules for the construction of any table, since this must be controlled largely by the purposes which the table is designed to meet. Data compiled for general use should be tabulated in such a form that the information presented is available to investigators from various angles, with equal facilities. In the preparation of tables for commercial use, the main object must always be to present the statistics in a manner which enables an executive to grasp their significance in a minimum amount of time, as otherwise they will probably prove useless for purposes of formulating commercial policies.

Since the manner of compilation of tables is controlled by the information to be tabulated and the purposes for which the tables are designed, it is possible to suggest only a few basic principles which must be kept in mind.

The tables may be simple or complex in character. A simple table is designed to give only one set of figures as in Table XIX, which seeks to show only the production of each of a series of machines.

TABLE XIX

PRODUCTION OF MACHINES FOR WEEK ENDING 5TH APRIL, 1944

Machine No.	Output
1	100,000
2	50,000
3	90,000
4	40,000
Total	280,000

It is rare, however, that such tables as the above are sufficient to convey all the information necessary, so that further columns must be included in order to incorporate additional information, vital to a proper consideration of the problem. In such a case as the production figures shown above, it would be necessary to include data to indicate not only the production, but also the amount of wastage. If it is desired to make comparisons, it would be necessary to include corresponding data for a previous period. To accomplish this it is necessary to add further columns to the table, which then becomes a complex table such as Table XX.

This table illustrates the value of incorporating comparisons in statistical compilations. In the first place, it is obvious that the production obtained from machines numbered 1 and 3 has improved as compared with the previous period. It must not be assumed without further investigation, however, that this is the result of increased efficiency, since the machines may

TABLE XX

PRODUCTION OF MACHINES FOR WEEK ENDING 5TH APRIL, 1944, AND CORRESPONDING PERIOD FOR 1943 TOGETHER WITH WASTAGE

	5TH APRIL, 1944			CORRESPONDING PERIOD FOR 1943		
(1) Machine No.	(2) Production	(3) Wastage	(4) Percentage of Wastage on Production	(5) Production	(6) Wastage	(7) Percentage of Wastage on Production
1	100,000	10,000	10·00	80,000	4,000	5.00
2	50,000	10,000	20·00	60,000	3,000	5·00
3	90,000	4,500	5·00	30,000	3,000	10·00
4	40,000	2,000	5·00	50,000	1,000	2·00
Total	280,000	26,500		220,000	11,000	

have been working longer hours. Should this be the case, it may be necessary to include additional columns to show the actual length of operating time in each period. Assuming, however, that the hours worked by each machine are equal and the same in each period, other conclusions may be made from the data included in the table.

It is immediately obvious that there is a very wide divergence between the output obtained from the individual machines. In the absence of other factors, such as the age of the machines, this is a question which calls for further investigation. Inquiries into this particular aspect of the data may reveal that certain machines are not being operated as efficiently as possible.

The table also illustrates the effect of incorporating percentages to show the ratio between various related figures. Thus, it would be almost impossible to see the significance of the wastage figures included in the table without the use of the percentages which have been included. These figures show very clearly that while there has been an improvement in the output obtained, this has also been accompanied by a more than proportionate increase in the figure for wastage in the course of production. This may be due to a speeding up of the operation and should cause the management to consider whether the value of increasing the production by this means is economical, or whether the loss through faulty work is so serious as to obviate any advantage which may accrue from the increased output. This may conceivably be the case when the raw material is of a very valuable character. Even if this is not the case, the table suggests that with greater care in carrying out the operations the proportion of scrap may be considerably reduced.

Finally, the table illustrates the fact that additional information may be tabulated by the inclusion of extra columns, with the result that more accurate and more detailed information is given. The number of columns actually included in any table, however, should be seriously considered, since it must be remembered that the more columns included in a table the more complex the table becomes, with the result that it is correspondingly more difficult to interpret. In some instances, it may be preferable to sub-divide the data into a number of independent tables, and finally to incorporate in one master table the really vital information contained in each of the subsidiary tables. When additional columns are included, the position of each column must also receive attention. This is essential, since the value of the tables will be enhanced by placing data which are recorded for purposes of comparison in adjacent columns, and, wherever possible, incorporating columns to show the exact relationship between the comparable data. In this way comparisons are immediately apparent, so that further calculations are rendered unnecessary in order to obtain the information which the table is intended to convey.

In compiling a table, only relevant information should be included. Thus, it is not usually of any value to incorporate data relating to a number of factors which are in no way related. Yet, if a relationship actually exists, considerable information is obtained by presenting all the data in one table. This is very clearly illustrated by Table XXI, which represents an analysis of the accounts of a company over a number of years.

This analysis compares the trend of turnover with the trend of costs, including material, wages, and factory expenses, and also the movement of the residual gross profit. In addition to including the total for each of the above items in the respective years, the value in respect of each article produced and sold has also been calculated. This method of analysis enables comparisons to be made of the value received for each article, and also of the various items of expenditure in the actual process of production. In the table under consideration it has been assumed that each article conforms to a universal standard, since, if this were not actually the position, the data would, first of all,

TABLE XXI

OF ACCOUNTS OF THE X CO., LTD

Year	(1) Sales	(2) Sales Value per Article Produced	(3) Material Cost	(4) Material Cost per Unit Produced	(5) Percentage on Sales	(6) Total Wages	(7) Wages Cost per Unit Produced	(8) Percentage on Sales	(9) Expenses	(10) Expenses per Unit Produced	(11) Percentage on Sales	(12) Gross Profit	(13) Gross Profit per Unit Produced	(14) Percentage on Sales
	£	s. d.	£	s. d.		£	s. d.		£	s. d.		£	s. d.	
1	100,000	10 –	50,000	5 –	50·0	30,000	3 –	30·0	10,000	1 –	10·0	10,000	1 –	10·0
2	90,000	12 –	40,000	5 4	44·4	25,000	3 4	27·8	15,000	2 –	16·7	10,000	1 4	11·1
3	120,000	8 –	75,000	5 –	62·5	30,000	2 –	25·0	10,000	8	8·3	5,000	4	4·2
4	150,000	7 6	90,000	4 6	60·0	40,000	2 –	26·6	8,000	4·8	5·4	12,000	7·2	8·0

have to be analysed between the respective classes of goods.

Assuming the homogeneity of the company's products, much valuable information can be obtained from the study of this analysis of the accounts. The table at once suggests that the turnover has increased over the period under consideration, but a glance at the figures of gross profit indicates that, although this increase in turnover has been achieved, the proportion of profits has not increased. In fact, the percentage on sales shows a decrease from 10 per cent to 8 per cent. This tendency did not apply until after the second year, since the gross profit for this year represented 11·1 per cent of the sales. The data presented in the table, however, supplies the necessary information to show why this apparent decline in results has occurred.

Column 2 represents the average price at which the products have been sold, and these figures are obtained by dividing the total sales in value by the total number of articles sold. Referring to these figures, it will be seen that in the second year the price obtained was considerably higher than in the first year. Thus it would appear that, although the actual turnover was not so great, the margin of profit was increased by reason of the better prices obtained for the goods sold. In subsequent years, the much-increased turnover seems to have been achieved by selling at lower prices, since the average value per article fell from 12s. to 8s. and 7s. 6d. respectively. In the third year, when the selling price was first reduced, the actual profits suffered considerably, but in the final year the increase in the profits suggests that the policy of selling at cheaper prices ultimately proved successful. Further information is obtainable from the data relating to the cost of production. It will be observed that in the third year the serious fall in profits was occasioned by a fall in the selling value of the article which was not accompanied by any appreciable reduction in the costs of production. Material used, while remaining at 5s. per article as in the first year, represented a higher proportion of the sales value. This is shown by the percentages included in Column 5, from which it is found that the material cost increased from 50 per cent to 62·5 per cent of the selling value. Savings in other directions, however, brought about a decrease in the proportions represented by wages and general expenses. The figures for the fourth year indicate that steps have been taken to reduce the material cost to a figure of 4s. 6d. per article, and also to effect further reductions in general expenses. At the same time, it should be carefully observed that wages have tended to increase in their proportion to sales. The information thus shown by the table suggests that this factor should be carefully considered in relation to future production.

Finally, the table again illustrates the value of including columns to show by means of percentages the actual relationship between the figures to be compared. In this case, the important factors to be considered are the respective proportions which the various classes of expenditure incidental to production bear to the total sales value, since any increase in these expenses must, of necessity, correspondingly decrease the residual proportion of gross profit.

A further consideration which must receive attention in compiling certain tables is the extent to which the actual figures are to be retained, and the extent to which the table can be simplified and improved by the adoption of approximations. Where the data consist of figures incorporating a large number of digits, it is usually confusing to include the whole of them, and if, therefore, this course is adopted, the table is difficult to examine, and comparisons become correspondingly more complicated. This fact is illustrated by Table XXII, which relates to data such as are included in the Bank Clearing returns. In this table the figures have been approximated to the nearest £1,000, and even when this has been done, it is still necessary to include data involving eight figures. It will be realized how unwieldy the table would become if the exact figures of pounds, shillings, and pence were included. While the table is thus rendered far more easy of interpretation by the adoption of approximation, it loses nothing of its value since the main features of increases and decreases are thereby much more clearly shown.

In conclusion, the title of the table in general and

TABLE XXII

BANK CLEARING RETURNS

	Town £000	Metropolitan £000	Country £000	Total £000
1st January to 6th March .	10,954,091	543,201	922,431	12,419,723
Week ending 13th March .	413,414	44,905	84,782	543,101
Week ending 20th March .	650,123	43,659	95,663	789,445
Total to date, 1934 . .	12,017,628	631,765	1,102,876	13,752,269
Total to date, 1933 . .	8,633,672	605,432	902,754	10,141,858
Increase (+) or decrease(–)	+ 3,383,956 39·2%	+ 26,333 4·3%	+ 200,122 22·2%	+ 3,610,411 35·6%

of each individual column must be stated so concisely that no ambiguity as to the real significance of the contents of the tables may arise. This is most important, since any doubt as to the true meaning of the facts presented would entirely destroy the value of any table, and may quite possibly result in erroneous conclusions being deduced therefrom.

EXERCISES

1. Draw up a skeleton table to show the earnings of men engaged in the cotton trade in ten separate towns, showing the numbers with earnings classified under each of the following groups: under 20s., 20s. and under 25s., 25s. and under 30s., 30s. and under 35s., over 35s. The table must show the total number employed, the average earnings for each town, and the percentage in each group.

2. Draw up a skeleton table in which could be shown the numbers employed in various industries at the time of the 1921 and 1931 censuses, distinguishing males and females, and showing also the numbers in six age groups, e.g. from the table it should be possible to discover the difference between the numbers of men from 25–35 years of age in the boot and shoe industry at the two dates.

CHAPTER IX

THE GRAPHIC METHOD

HOWEVER clearly tables may be compiled, a certain amount of time is inevitably necessary to ascertain the salient features presented. In fact, in many cases it is difficult for anyone other than a trained mathematician to grasp the full meaning of data set down in tabular form except after very careful study. The business man, however, must have facts presented in such a manner that the important features are immediately apparent, and, for this reason, the graphic method has become increasingly important in the presentation of facts relating to business problems. It is possible to present almost all classes of data in graphic form, and the speed and ease with which the principal facts can be grasped by this method have resulted in its increased adoption in commercial statistics.

In order to illustrate the method of presentation of data, consider the following Table XXIII, which indicates the monthly sales and purchases of a company over a period of two years. From a study of this table very little information can be obtained. It is possible to see that purchases rose until May, and that the sales fluctuated between £45,000 and £68,000 per month. Beyond this it is impossible to go. It is impossible to delineate any relationship between the governing policy as applied to purchases and sales or determine any relationship which exists between the

TABLE XXIII

TABLE SHOWING PURCHASES AND SALES (MONTHLY) FOR ONE YEAR

Month	Purchases	Sales
	£	£
January	50,000	60,000
February	60,000	68,000
March	65,000	50,000
April	65,000	45,000
May	70,000	46,000
June	60,000	50,000
July	40,000	55,000
August	40,000	62,000
September	50,000	65,000
October	50,000	65,000
November	54,000	60,000
December	50,000	63,000
Total	£654,000	£689,000

two sets of data. The real importance of these figures is, therefore, scarcely realized until they are shown graphically as in Diagram 3.

A glance at the graph drawn to represent this data shows in striking manner the course of purchases and sales during the year. The curve of the purchases indicates immediately the particularly heavy purchases during the first five months of the year. This was actually due to the fact that stocks had been allowed to fall at the end of the previous year, so that the policy governing the purchases over these months was controlled by the necessity for building up the stock. Secondly, the purchases were heavy in anticipation of sales, which previously had tended to increase after the first few months of each year. A complete reversal of this tendency is seen to have taken place after May, a considerable drop being recorded for the months of June and July. This alteration in policy was the result of two considerations. Firstly, the sales for the period from January to May had fallen far short of expectations. This, therefore, immediately indicated that as the purchases were so heavy, liabilities were being incurred in respect of these purchases to the extent of £310,000 to the end of May, whereas the income from sales could only be £269,000. This suggested that if the policy of making large purchases were continued, the position would quickly become difficult owing to liabilities exceeding income by so great a margin.

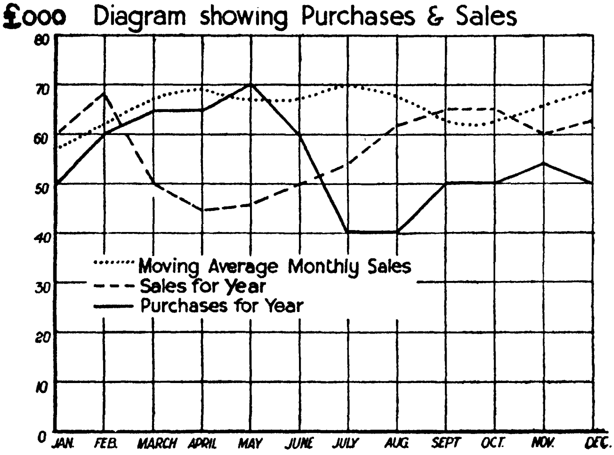


DIAGRAM 3

A second consideration also affected this change in policy. A study of market conditions in which the concern was interested, together with internal information available, indicated that an appreciable reduction of the market prices of the goods was about to take place. It was, therefore, obviously not desirable that heavy stocks should be on hand when this reduction took place as this would have involved the company in considerable losses in respect of excess stock carried. For the remaining portion of the year purchases continued to be low, governed entirely by the magnitude of sales, goods being ordered only to enable turnover to be maintained, thus retaining stocks at as low a figure as was consistent with sales.

Turning now to a consideration of the facts regarding sales, we immediately see that the turnover fell seriously from February until May, and from then onwards showed a gradual increase until the end of the year. This may have been an ordinary movement for sales, which may have been of a seasonal nature normally showing a falling off for the months from February to July. This, however, must not be assumed without further investigation. In carrying out such an investigation it is not sufficient to compare the sales figures for the year under review with those for the previous year, since exceptional conditions may have prevailed in either period. A true conception can only be obtained by comparing the year's figures with a moving average calculated over a number of years, thereby eliminating exceptional items. In order that such a comparison can be made, a five-year moving average has also been included in the graph. To obtain the necessary data to plot this curve the average sales for January in each of five years are calculated and similar averages obtained for each month, so that a moving average is then obtained for the twelve months. By plotting these figures on the graph it is seen that seasonal fluctuations are unusual, and that the sales show a tendency to remain fairly even over the entire year, fluctuating only between narrow limits. The diagram, therefore, clearly indicates that the falling off in sales recorded for the year under review is exceptional, and further analysis is necessary to ascertain the exact cause. In actual fact, it was found that the falling off was due to the entrance of serious additional competition in the market, combined with a tendency for prices to decline. The position being realized at an early date, increased efforts, coupled with extensive advertising, enabled an improvement to be achieved, so that before the end of the year sales regained more normal figures.

In the foregoing illustration, the facts have been considered after the year has been completed, but, naturally, such diagrams would be continually brought up to date so that any exceptional features would immediately be revealed and steps taken to remedy any adverse conditions.

It is frequently advantageous to indicate the trend of sales or other data without conveying exact figures, and it follows that it is impossible to achieve this object by means of tabulation. This, however, can be done by means of graphs. As an example, Diagrams 4, 5, and 6 are similar to those which are found in actual use, and are designed to show each salesman the progress he is making during the year. The sales anticipated from each salesman for each month are decided at the commencement of each year, and are immediately plotted on the graph relative to each

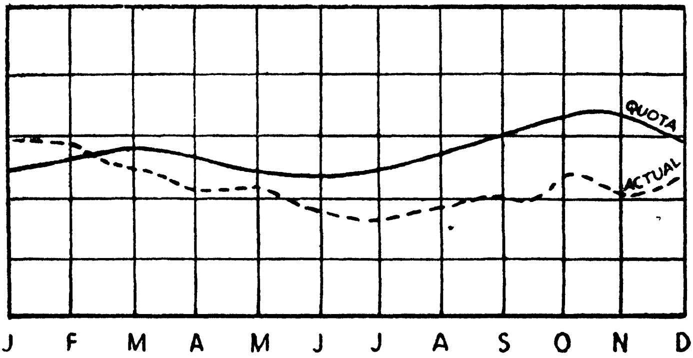


Diagram 4. Salesman 1

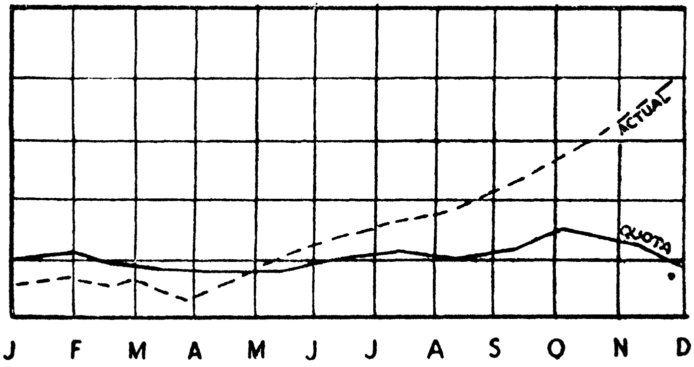


Diagram 5. Salesman 2

salesman. In fixing the individual quotas, due consideration is given to the normal seasonal fluctuations in sales. Against this curve the actual current sales are plotted, so that each salesman can see immediately to what extent he is achieving his quota or to what extent his actual sales are falling short of the turnover expected from him. It is not desirable that the actual figures should be revealed to every salesman, since lower quotas may be fixed for representatives working a particularly difficult area, and also it is desirable that they should be of no significance to individuals unconnected with the concern who may see the diagrams. For this reason the magnitudes have been omitted from the graphs, which, although serving the purpose of showing the success which each representative is obtaining, do not divulge valuable private information. The three graphs show very clearly, and without the necessity for any calculation, the following details—

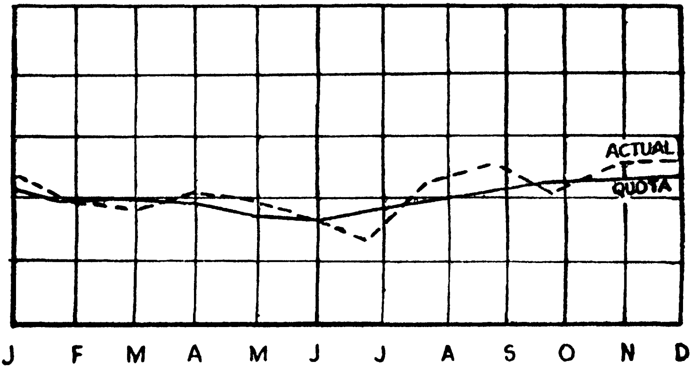


DIAGRAM 6. SALESMAN 3

1. The first salesman was a recognized representative from whom much was expected. In view of the fact

that he was employed in an area which had already been extensively developed, a high quota was fixed. The diagram shows that this salesman seriously failed to achieve the required turnover, indicating that circumstances are operating which require immediate investigation.

2. The second salesman was engaged in a new area in which sales had not previously been cultivated, but the graph testifies to the excellency of his efforts, since he has considerably exceeded the turnover anticipated.

3. Finally, it will be noted from the graph relative to the third salesman that this man has just covered his quota in the majority of months, and in one or two cases has just failed.

Turnover and sales are by no means the only problems in which the graphic method can be employed to present data. The method may be employed with equal success to almost any problem. For example, it is possible to show graphically the weekly range of wages, production at various periods, and by plotting figures for one year against those for another year or against a moving average, valuable comparisons can be made. Additional information is also obtained if on the same graph statistics of wages and output are incorporated. Such a graph would immediately reveal when, for example, an increase in wages was not synchronous with an increase in output. Similarly, it may be found that the sales represented by quantities may increase considerably, whereas the monetary value of these sales, owing to falling selling prices, may show only a slight increase. In order to present these facts in the most striking manner, it is only necessary to plot on the same graph data of monetary sales and quantity sales as in Diagram 7.

Logarithmic Graphs

The graphs so far discussed have been based on what is called the "Natural Scale." In some cases, however, the natural scale graph is not a suitable medium of presentation, since it is possible to show only the actual magnitude of increases or decreases by this method, and not the rate of change. For example, in considering the growth of trade it is necessary to indicate the rate of increase or decrease rather than the actual magnitude of the fluctuation. This may be illustrated by reference

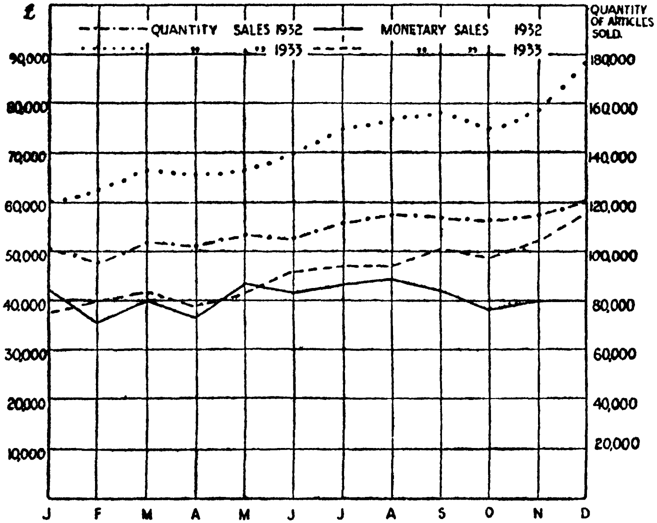


DIAGRAM 7

to the following data relating to the London Bank Clearings for four months—

	London Clearings	Increase
January . . .	1,000,000,000	
February . . .	1,200,000,000	20%
March . . .	1,440,000,000	20%
April . . .	1,728,000,000	20%

The data shown above would be represented on the normal scale as in Diagram 8, which shows both the level and the amount of the movement of the London Clearings, but does not indicate the rate of change which has taken place from month to month. The series accordingly slopes upwards with increasing sharpness, and may thus lead to the conclusion that growth was continually augmented. From the data, however, it is known that the figures for each month

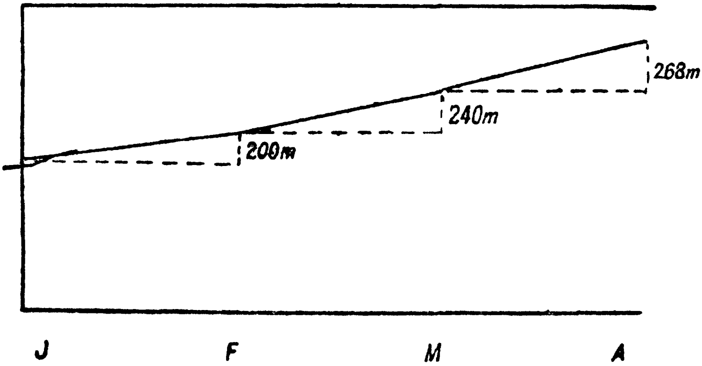


DIAGRAM 8

represented a constant increase of 20 per cent on the previous month. In considering the development shown by such statistics, it is more important to show the rate of growth than the magnitude of the increase, and to do this the logarithmic or ratio scale must be adopted. Under this method, equal proportionate advances are represented by equal rises in the slope of the series as shown in Diagram 9.

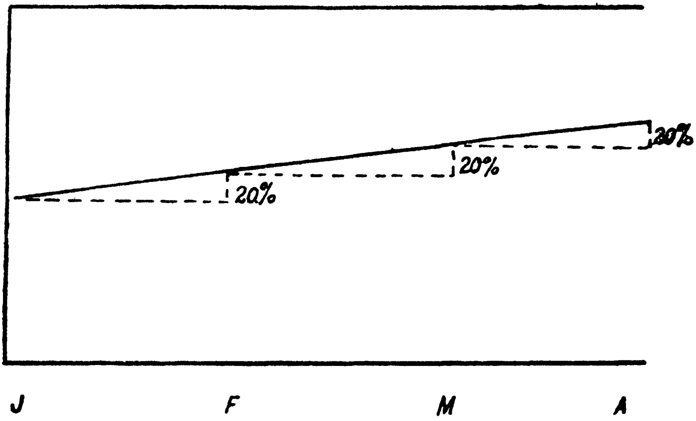


DIAGRAM 9

In constructing a logarithmic graph, the data are converted into logarithms, and the curve is plotted from the logarithms instead of from the actual figures. In this manner the graph is so constructed that equal percentage increases or decreases in periods, however remote, are represented by equal movements of the curve, while equal actual increases or decreases are reduced to their respective ratios. Care must be taken to remember that in a logarithmic graph ratios of rise and fall are shown and not the rises and falls themselves. For this reason, it is frequently necessary to show side by side the effect of presenting the data in each form. It will be obvious that the logarithmic graph should be employed only when it is desired to show ratios. This may be the case when it is intended to show the gradual development of trade or of a particular business, since, in this way, means are provided whereby the increase or decrease in the rate of growth may be studied.

To illustrate the commercial application of logarithmic graphs, the following data of gross profits and wage payments have been presented on the natural scale and on the logarithmic scale in Diagrams 10 and 11.

TABLE XXIV

GROSS PROFITS AND WAGES TO NEAREST £1,000

Year	1919	1920	1921	1922	1923	1924	1925	1926
Gross profits .	205	310	170	170	185	200	210	170
Wages . .	120	170	112	120	130	145	150	135

In the logarithmic graph, it will be observed that no zero line must be shown, as otherwise this would almost certainly lead to erroneous conclusions. Of the two diagrams, the logarithmic graph is the more valuable, since the proportionate increases are more important in following the rates of development than the actual increases. Accordingly, more valuable information is given by showing that in 1920 profits increased by 51·2 per cent and wages by 41·7 per cent, than by showing that profits increased by £105,000, compared with an increase of £50,000 in wages.

Conclusion

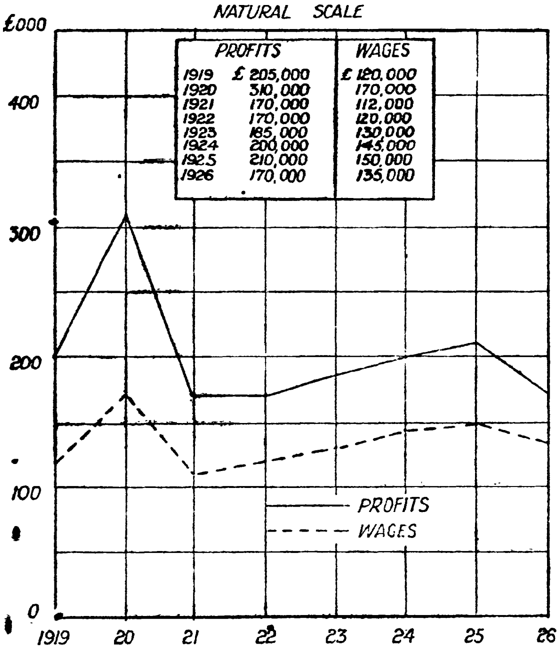
Sufficient has been said to indicate that the graphic method is adaptable to any commercial problem. Financial data, advertising results, the relationship between material, labour, and other costs of production, factory records of output, and turnover records can all be shown far more successfully by means of graphs than by any other method. After careful consideration of the problem involved, suitable methods of graphic presentation can be decided upon, and logarithmic scales adopted where necessary. By studying the curves so obtained and the tendencies indicated, it is possible to make a forecast of the probable future tendencies of trade conditions. This principle has, for example, been applied by the London School of Economics and the Harvard University Committee in studying the movements of curves of speculation, business, and money in order to ascertain the relationship between the movements of these factors with a view to forecasting the possible development of trade in England and America respectively.

Finally, the use of the graphic method is not limited by complexity in compilation, since no great skill is required either to construct or to interpret a graph.

EXERCISES

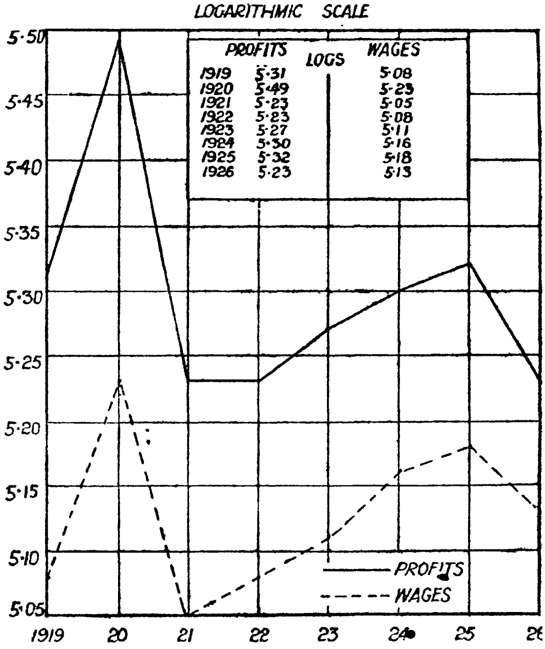
1. Draw a graph showing the monthly movements of the following data. Calculate the monthly moving averages, and

Profits and Wages on Natural and Logarithmic Scale



	PROFITS	WAGES
1919	£ 205,000	£ 120,000
1920	310,000	170,000
1921	170,000	112,000
1922	170,000	120,000
1923	185,000	130,000
1924	200,000	145,000
1925	210,000	150,000
1926	170,000	135,000

Diagram 10



	PROFITS	LOGS	WAGES
1919	5·31		5·08
1920	5·49		5·23
1921	5·23		5·05
1922	5·23		5·08
1923	5·27		5·11
1924	5·30		5·16
1925	5·32		5·18
1926	5·23		5·13

Diagram 11

alongside the graph show the moving average figures obtained. Comment on any facts indicated by the graph—

	Jan.	Feb.	Mar.	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
1924	203	150	110	110	100	60	50	50	80	210	250	350
1925	244	150	140	140	200	150	100	100	80	150	240	270
1926	256	177	180	120	110	50	70	40	40	160	300	290
1927	330	290	250	200	150	90	70	40	50	150	240	280
1928	310	260	140	70	70	80	70	43	40	150	230	370
1929	340	296	170	90	120	110	70	50	40	180	270	220
1930	145	120	100	120	100	86	50	50	70	220	330	390
1931	300	227	150	80	70	80	50	40	90	250	300	300
1932	329	266	270	220	150	60	50	50	80	250	380	400
1933	300	188	120	100	130	80	60	40	80	230	310	300

2. Draw a graph to show the comparison between the output of three factories as shown in the following data—

	Factory 1	Factory 2	Factory 3
1928	191,000	127,000	50,000
1929	211,250	101,000	65,400
1930	173,450	104,000	56,500
1931	142,700	70,000	75,800
1932	116,900	90,000	95,400
1933	101,400	100,000	150,700

CHAPTER X

DIAGRAMS

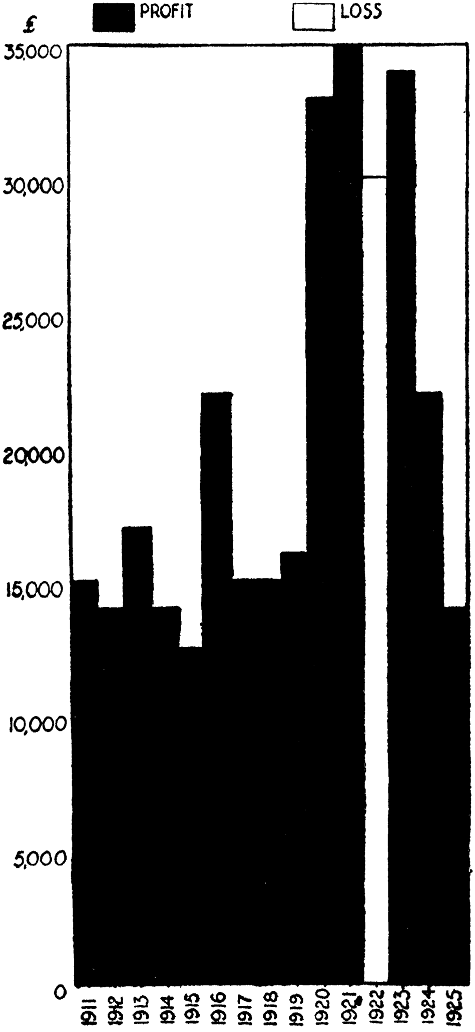
TABULATION and the graphic method provide the most important and the most satisfactory means of presenting statistical data. Occasions arise, however, when it is necessary to show the facts in such a manner that those who are not conversant with either a table or a graph can readily see the significance of the figures. It is mainly for such purposes that block and similar diagrams are resorted to, their usefulness being restricted to the presentation of fairly simple classes of data.

An example of the diagrams to which reference has been made may be considered. These diagrams show alternative methods of presenting the figures of earnings of a certain company, the data used being shown in the following table—

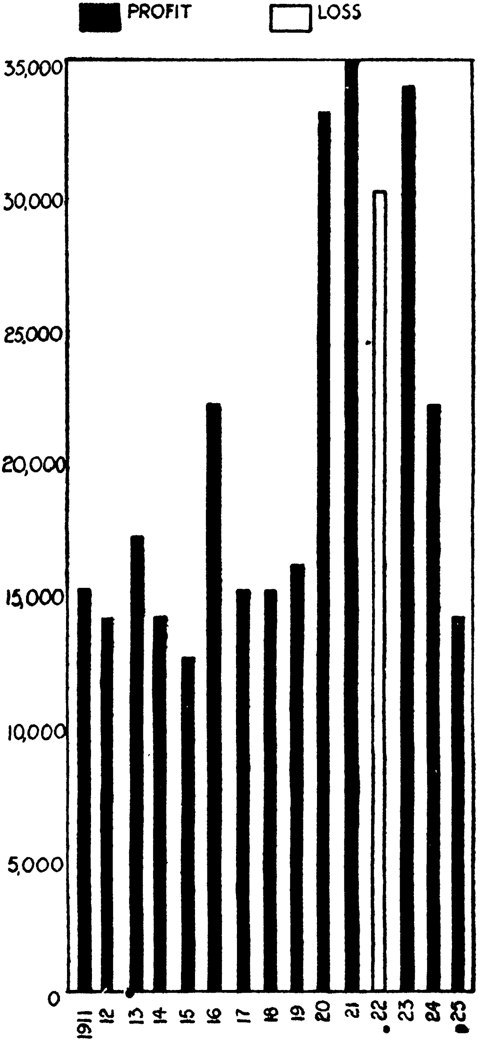
TABLE XXV

TABLE OF PROFITS OF THE X CO., LTD.

Year	Profit	Loss	Year	Profit	Loss
	£	£		£	£
1911	15,000		1919	16,000	
1912	14,000		1920	33,000	
1913	17,000		1921	35,000	
1914	14,000		1922		30,000
1915	12,500		1923	34,000	
1916	22,000		1924	22,000	
1917	15,000		1925	14,000	
1918	15,000				



Block Diagram 12



BAR DIAGRAM 13

The examples show a comparison between a block diagram and a bar diagram. It will be observed that the block diagram is more difficult to read than the bar diagram, and it will also be realized that it would

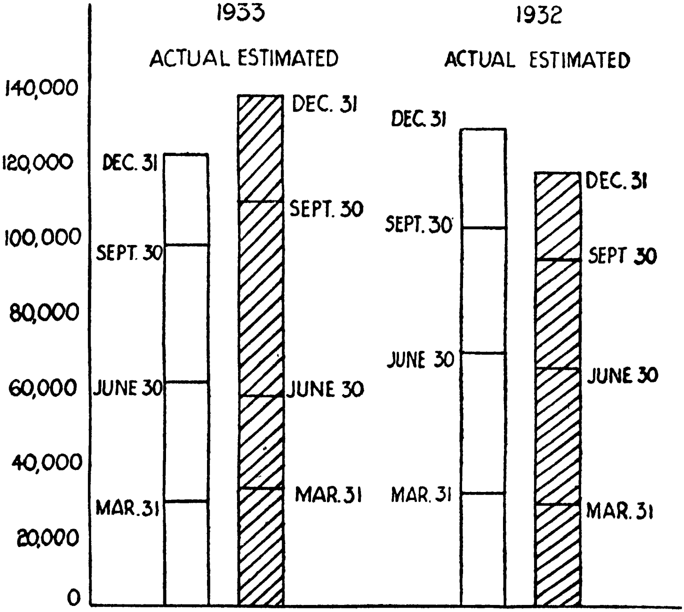


Diagram 14. Comparison of Estimated and Actual Expenditure 1933 and 1932

be difficult to make comparisons between the earning capacity of one concern and another under this method. This is due to the fact that it is impossible to depict the earnings of more than one company on one diagram, and, if it is desired to make such comparisons, it is possible to do this only by drawing two diagrams and placing them side by side. If the data had been presented on a graph, any reasonable number of comparisons could be made by plotting on the same graph separate curves for each set of data.

The pictogram is a variation of the bar diagram, and is illustrated by Diagram 14, which is designed to show the estimated and actual expenditure incurred in two separate years at various dates. Diagram 15 is a similar diagram showing the analysis of the actual expenditure

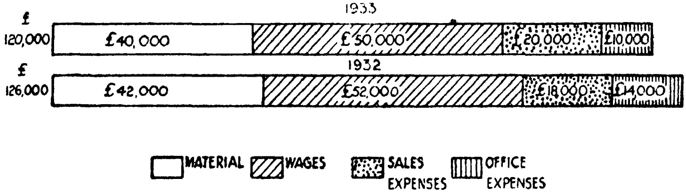


DIAGRAM 15. ANALYSIS OF ACTUAL EXPENDITURE, 1933 AND 1932

for each year. In the first case, the scale on which the diagram is drawn is shown on the left-hand side to facilitate reference, and the method thus enables the data to be presented in a striking manner. The pictogram is complicated, however, by the fact that only the first section in each diagram starts at the same point. To show up each part plainly, each section is coloured differently, or alternatively different types of shading are employed. In addition, each diagram must be reasonably narrow in order to emphasize the fact that only linear measurements are to be considered, and that the area does not enter into the comparisons. This type of diagram does not lend itself to the illustration of any other than relatively simple data.

It is frequently advantageous to express additional facts by means of incorporating the area of a diagram in the comparison. Thus, in Diagram 16, additional information is given by this method in respect of the earnings and expenditure of two workmen. In each case, the length of the rectangle represents the number of hours worked by each workman, and the width is employed to illustrate the respective hourly rates. Consequently, the areas of the rectangles represent the relative total earnings of the two men. By subdividing the total area in each case, a comparison is given of the proportionate expenditure on food, rent, clothes, and other items respectively.

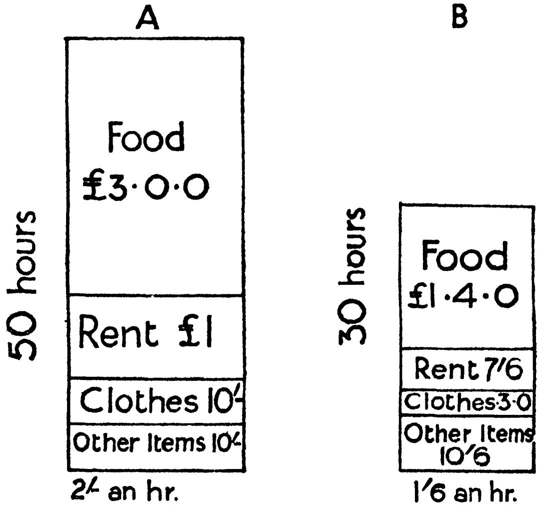


DIAGRAM 16. DIAGRAM SHOWING COMPARISON OF EARNINGS AND EXPENDITURE OF A AND B FOR ONE WEEK

Such diagrams are frequently employed to convey to laymen more or less technical data which it is desired that they should understand. Thus, to the average citizen, the Abstract of Accounts of a County Borough conveys very little information. If, however, the data is presented diagrammatically, the important facts can be conveyed to him. This has been done in Diagram 17. The total area of the circle represents the

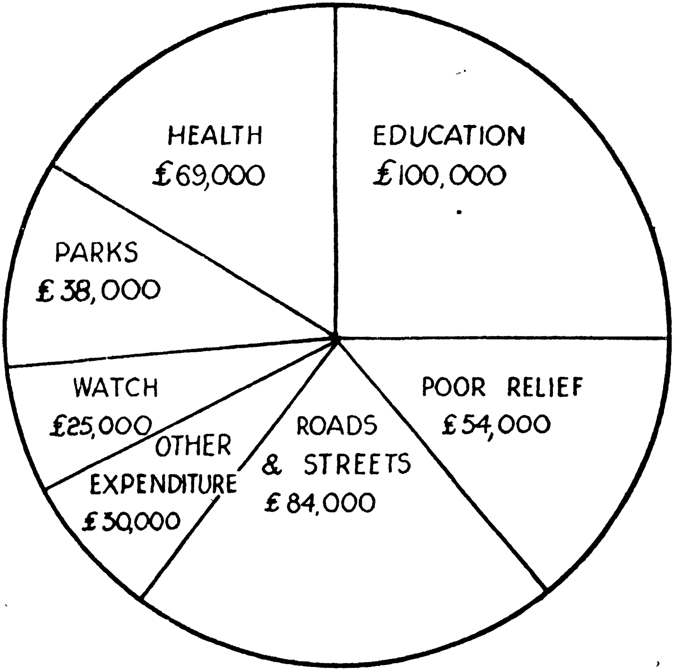


DIAGRAM 17. SHOWING ALLOCATION OF INCOME FROM RATES FOR A COUNTY BOROUGH, TOTAL RATE £400,000

entire income from rates, and each segment, drawn proportionately, shows the allocation of the rates to various classes of expenditure. The diagram shows at a glance the proportion of the rates which is spent on each particular service. It also emphasizes various important facts. The heavy expenditure in respect of health, education, and maintenance of roads and streets is clearly indicated. The Borough also spends on Relief of the Poor more than half of the expenditure on Education and nearly one-seventh of the total income from rates, while the expenditure on Police, etc., is rather smaller than might be expected.

It will be observed that diagrams of one or other of the above types are frequently used for advertising purposes. This is possible because the methods enable facts relative to various products to be shown to the prospective consumer in a striking manner without the assumption of any knowledge of statistics. Probably one of the most effective applications is in the case of prepared foods in which the relative proportions of vitamins or other constituents included in the advertised products are compared with those in competitive products.

In the presentation of data in this manner, the utmost care must be taken to ensure that erroneous conclusions are not made from the diagrams. Confusion may very readily arise where data are represented by areas. Consequently, it is of particular importance that the manner in which the diagrams have been designed should be concisely stated. For example, comparisons may be made by representing two sets of data by relative circles. In such cases it is possible to draw the circles either with proportionate areas or with proportionate diameters. The difficulty of comparisons between circles drawn on either method, however, precludes the possibility of accurate conception of their true values. The practical use of circles should, therefore, be mainly confined to the representation by means of segments of the varying constituents of the whole, according to their proportionate value.

Diagrams of this nature are not found to be entirely satisfactory for the presentation of statistical data compiled for the purpose of solving preconceived problems, but they may conveniently be used when it is desired to present in a striking manner relatively simple data for the information of those not acquainted with technical methods.

EXERCISES

1. Draw a diagram to represent the distribution of employees of a factory according to age as shown in the following table—

Over	15–20	100
,,	20–25	90
,,	25–30	1,000
,,	30–35	500
,,	35–40	200
,,	40–45	200
,,	45–50	250
,,	50–55	50
,,	55	10

2. Suggest a diagram which would be suitable for the publication of a comparison between the rate at which income tax was collected by the Inland Revenue in two comparative years. The diagram should also provide for a comparision between the actual receipts and the annual estimates for each period.

CHAPTER XI

CORRELATION

In the consideration of the significance of any series of data, the average, the median, or the mode may be employed to give a summary of the data which can be used for comparison purposes. Within any group, however, the items will vary from one another and from any type which is employed owing to the differences in their magnitude. Again it may be found that in comparing two sets of data, fluctuations in one group may be accompanied by similar proportionate fluctuations in another group, or alternatively the fluctuations in the first group may be followed by inverse fluctuations in the second.

It will be seen, therefore, that in studying any data, consideration must be given to the deviation of the items from the type employed, and, secondly, that where comparisons are to be made between two or more sets of data, the extent to which they are in any way related must also be investigated. Such a relationship, when it exists either directly or indirectly, is known as correlation.

Deviation

It is first of all essential to obtain a summarized expression of the fluctuations which occur in the individual items of a group, and it must be remembered that the methods of comparison which are discussed are designed to give a numerical indication of these fluctuations.

In Table XXVI, Columns 1 and 2, show the output of factories A and B for the past year. It will be noted in comparing these two sets of data that the annual output, and therefore the average monthly output, is the same for both factories. It will be observed from a further study of the figures, however, that the monthly output figures for factory A show much greater fluctuations from the average than in the case of factory B. The simplest method of considering this deviation would be to consider the range between the lowest and greatest magnitude. Thus for factory A the range is from a magnitude of 60 to 140, and in the case of factory B from 90 to 110.

This method, however, is obviously unsuitable for the required purposes, since it depends largely upon the items which actually happen to have been taken, which is an important consideration when the data have been compiled by sampling, and, in addition, the method does not lend itself to mathematical treatment.

Average Deviation

The average deviation is obtained by calculating the deviation for each item, and multiplying by the magnitude of the item. The total of these products is then divided by the total magnitude of the items to obtain the average deviation, as illustrated on page 128.

It is, therefore, necessary to obtain a numerical expression indicating the extent of the deviations in the individual items. There are two principal methods employed in the measurement of deviation—average deviation and standard deviation.

TABLE XXVI

SHOWING MONTHLY OUTPUT OF FACTORIES A AND B AND CALCULATION OF AVERAGE DEVIATION

	Factory A. Output *f*.1	Factory B. Output *f*.2	Deviation from Average. Factory A *d*.1	Deviation from Average. Factory B. *d*.2	Product of $f1 \times d1$	Product of $f2 \times d2$
Jan. . .	90	100	10	—	900	—
Feb. . .	110	102	10	2	1,100	204
Mar. . .	120	105	20	5	2,400	525
Apl. . .	110	110	10	10	1,100	1,100
May . .	120	105	20	5	2,400	525
June. .	140	103	40	3	5,600	309
July . .	90	95	10	5	900	475
Aug. . .	110	107	10	7	1,100	749
Sept.. .	60	90	40	10	2,400	900
Oct. . .	80	90	20	10	1,600	900
Nov.. .	90	95	10	5	900	475
Dec. . .	80	98	20	2	1,600	196
Total .	1,200	1,200	—	—	22,000	6,358

Average deviation—

Factory A . . $\dfrac{22,000}{1,200} = 18{\cdot}33$

Factory B . . $\dfrac{6,358}{1,200} = 5{\cdot}30$

By expressing the average deviation as a proportion of the average, the Average Coefficient of Dispersion is obtained, thus—

Factory A . . $\dfrac{18{\cdot}33}{100} = {\cdot}1830$

Factory B . . $\dfrac{5{\cdot}30}{100} = {\cdot}0530$

It will be observed that in interpreting the importance of the Average Coefficient of Dispersion the greater the coefficient the greater the extent of the fluctuations which take place in the individual items. Therefore in the example above the two coefficients indicate that the fluctuation in the case of factory A is much greater than in the case of factory B, which, of course, in such a simple case can be verified by inspection.

Standard Deviation

In calculating the average deviation it will have been noted that no distinction is made between those items which are above and those which are below the average, and that therefore some deviations should be shown as + and some as –. This objection is eliminated in the case of the second method of measuring the extent of the fluctuations, which is known as the standard deviation and forms the basis for carrying out further investigations into the question of correlation.

To obtain the measurement of standard deviation, the deviation shown by each individual item is squared, so that all signs become positive, and therefore the objection to the method employed in calculating average deviation is immediately disposed of. The squares so obtained are then totalled, divided by their number, and the square root of the figure so obtained represents the standard deviation. This method is illustrated in Table XXVII, in which case the standard deviation has been calculated for the same data as is employed in the previous illustration of the average deviation.

TABLE XXVII

SHOWING OUTPUT OF FACTORIES A AND B AND CALCULATION OF STANDARD DEVIATION

	Factory A. Output *f*.1	Factory B. Output *f*.2	Factory A. Deviation *d*.1	Factory B. Deviation *d*.2	Square of *d*.1	Square of *d*.2
Jan. . .	90	100	− 10	—	100	—
Feb. . .	110	102	+ 10	+ 2	100	4
Mar. . .	120	105	+ 20	+ 5	400	25
Apl. . .	110	110	+ 10	+ 10	100	100
May . .	120	105	+ 20	+ 5	400	25
June. .	140	103	+ 40	+ 3	1,600	9
July . .	90	95	− 10	− 5	100	25
Aug.. .	110	107	+ 10	+ 7	100	49
Sept.. .	60	90	− 40	− 10	1,600	100
Oct. . .	80	90	− 20	− 10	400	100
Nov.. .	90	95	− 10	− 5	100	25
Dec. . .	80	98	− 20	− 2	400	4
Total .	1,200	1,200	—	—	5,400	466

Standard Deviation, Factory A $\sqrt{\frac{5400}{12}} = 21{\cdot}21$

Standard Coefficient of Dispersion $\frac{21{\cdot}21}{100} = {\cdot}2121$

Standard Deviation, Factory B $\sqrt{\frac{466}{12}} = 6{\cdot}23$

Standard Coefficient of Dispersion $\frac{6{\cdot}23}{100} = {\cdot}0623$

The calculation of standard deviation shown above indicates that similar conclusions may be drawn as in the case of average deviation, but the method of standard deviation only can be employed when investigations into correlation are carried out.

Average deviation and standard deviation are in some cases calculated from the median, as an alternative to the deviation from the arithmetic average.

Correlation

In dealing with deviation, consideration has only been given to the question of obtaining numerical measures of one particular characteristic of a group of data.

Correlation, however, deals with the comparison between two or more different groups of data in order to ascertain whether any relation exists between them, and to obtain a numerical expression of the degree of such relationship.

Suppose now that the following table represents the capital invested in a company during the first ten years of its existence and the profits earned during this period. Since it is known that, eliminating exceptional circumstances, profits should increase according to the increase in the amount of capital employed, it is reasonable to assume that there should be some relationship between the two sets of data.

TABLE XXVIII

Year	Capital	Profits
	£	£
1	1,000	300
2	2,000	400
3	3,000	800
4	4,000	800
5	5,000	1,000
6	6,000	1,500
7	7,000	1,500
8	8,000	2,500
9	9,000	2,200
10	10,000	3,000

The existence of such a relationship between two sets of data can be clearly shown graphically. Care must be taken that a correct comparison is shown and in plotting the graph drawn in Diagram 18, which is a graphic presentation of the data referred to, this has been ensured by converting the figures into indices on the basis of the first year, and plotting a three-year moving average in order that short-term fluctuations may be eliminated. The figures are shown in Table XXIX.

TABLE XXIX

	CAPITAL			PROFITS		
Year	Actual	Index	Three-year Moving Average	Actual	Index	Three-year Moving Average
	£			£		
1	1,000	100	—	300	100	—
2	2,000	200	200	400	133	166
3	3,000	300	300	800	266	222
4	4,000	400	400	800	266	288
5	5,000	500	500	1,000	333	366
6	6,000	600	600	1,500	500	444
7	7,000	700	700	1,500	500	611
8	8,000	800	800	2,500	833	688
9	9,000	900	900	2,200	733	855
10	10,000	1,000	—	3,000	1,000	—

From a study of the graph it will be seen that as the curve of profits does not exactly follow the curve for capital, the relationship is not absolutely perfect, and therefore some numerical expression must be obtained to represent the extent to which the relationship does actually exist. It has been found that the most satisfactory method of obtaining such a measure is as follows: The deviation for each item in the first series is multiplied by the deviation for the corresponding item in the second series. The sum of these products is then divided by the number of items, multiplied by the

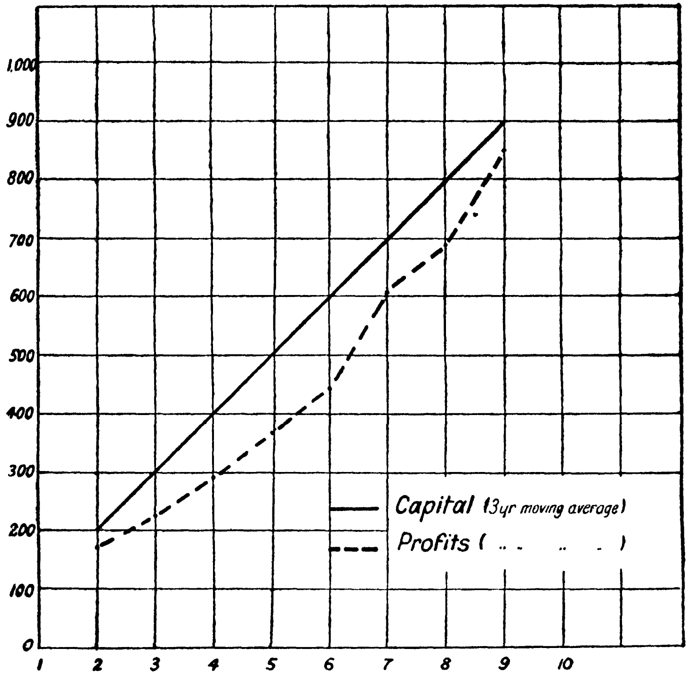


DIAGRAM 18

standard deviation for each series. This may be expressed by the formula—

$$\text{Coefficient of Correlation} = \frac{\Sigma\,(xy)}{n \times d_1 \times d_2}$$

where $\Sigma\,(xy)$ represents the sum obtained of the products of multiplying the deviations of the respective items in each group, n = the number of items in each group, d_1 = the standard deviation of the first group, and d_2 = the standard deviation of the second group.

In order to illustrate this principle, the formula may be applied to the example given above.

The standard deviation must first be calculated as in Table XXX. It will be noted that in the whole calculation the last two noughts have been eliminated as this does not affect the final result.

TABLE XXX

CALCULATION OF STANDARD DEVIATION

CAPITAL				PROFIT		
Year	1st Series £00	Deviation from Average	Square of Deviation	2nd Series	Deviation from Average	Square of Deviation
1	10	− 45	2,025	3	− 11	121
2	20	− 35	1,225	4	− 10	100
3	30	− 25	625	8	− 6	36
4	40	− 15	225	8	− 6	36
5	50	− 5	25	10	− 4	16
6	60	+ 5	25	15	+ 1	1
7	70	+ 15	225	15	+ 1	1
8	80	+ 25	625	25	+ 11	121
9	90	+ 35	1,225	22	+ 8	64
10	100	+ 45	2,025	30	+ 16	256
Total	550	—	8,250	140	—	752

$$\text{Standard Deviation, First Series } \sqrt{\frac{8250}{10}} = 28{\cdot}72$$

$$\text{Standard Deviation, Second Series } \sqrt{\frac{752}{10}} = 8{\cdot}67$$

The Calculation of the Coefficient of Correlation is then as follows—

Year	Capital, 1st Series £00	Deviation from Average (x)	Profits, 2nd Series £00	Deviation from Average (y)	Product of Deviations (xy)
1	10	− 45	3	− 11	+ 495
2	20	− 35	4	− 10	+ 350
3	30	− 25	8	− 6	+ 150
4	40	− 15	8	− 6	+ 90
5	50	− 5	10	− 4	+ 20
6	60	+ 5	15	+ 1	+ 5
7	70	+ 15	15	+ 1	+ 15
8	80	+ 25	25	+ 11	+ 275
9	90	+ 35	22	+ 8	+ 280
10	100	+ 45	30	+ 16	+ 720
					2,400

Coefficient of Correlation

$$\frac{2400}{10 \times 28{\cdot}72 \times 8{\cdot}67} = \frac{2400}{2490} = {\cdot}96$$

In interpreting the significance of a coefficient of correlation if the probable error is small and may be neglected, then when the coefficient is less than ·30 no correlation can be said to exist, but if it is ·50 or over, decided correlation may be assumed. Thus it may be inferred from the coefficient calculated for the data given above, that although the relationship is not perfect, the correlation is very great. It should be remembered that the coefficient will never be greater than 1·0, which represents perfect correlation; 0, represents absolute freedom from any relationship whatsoever, whilst − 1·0 represents perfect inverse relationship.

As an illustration of inverse relationship the question of the supply and price of a commodity may be considered, since it is known that where supply increases prices tend to fall, and *vice versa*.

Table XXXI therefore shows the indices of supply and price for a certain commodity, and the calculation of the coefficient of correlation between the two sets of data. It will be observed that a very high rate of inverse correlation exists, indicating that any increase in supply is almost invariably accompanied by a fall in price, and a fall in supply accompanied by an increase in price.

TABLE XXXI

CALCULATION OF COEFFICIENT OF CORRELATION BETWEEN INDEX NUMBERS OF SUPPLY AND PRICE

SUPPLY				PRICE			
Year	Index No.	Deviation from Average (x)	Square of Deviation	Index No.	Deviation from Average (y)	Square of Deviation	Product of Deviation (xy)
1	80	− 11	121	140	+ 25	625	− 275
2	80	− 11	121	141	+ 26	676	− 286
3	85	− 6	36	130	+ 15	225	− 90
4	90	− 1	1	120	+ 5	25	− 5
5	82	− 9	81	130	+ 15	225	− 135
6	84	− 7	49	126	+ 11	121	− 77
7	90	− 1	1	115	—	—	—
8	95	+ 4	16	94	− 21	441	− 84
9	104	+ 13	169	64	− 51	2,601	− 663
10	120	+ 29	841	90	− 25	625	− 725
—	910	—	1,436	1,150	—	5,564	−2,340

Standard Deviation, Supply $\sqrt{\frac{1436}{10}} = 11{\cdot}98$

Standard Deviation, Price $\sqrt{\frac{5564}{10}} = 23{\cdot}59$

$$\text{Coefficient of Correlation} = \frac{-2340}{10 \times 11{\cdot}98 \times 23{\cdot}59}$$

$$= -{\cdot}83$$

EXERCISES

1. Calculate the coefficient of correlation for the following data relating to the receipts and expenses per car mile for a transport undertaking—

Receipts: 9d. 9d. 8d. 7d. 6d. 7d. 10d. 11d. 13d. 10d.
Expenses: 6d. 7d. 5d. 4d. 3d. 6d. 7d. 7d. 8d. 7d.

2. Calculate the average deviation for each of the above series from

(*a*) Arithmetic average. (*b*) Median.

CHAPTER XII

BUDGETARY CONTROL IN BUSINESS

BUDGETARY control of business organizations is perhaps one of the most recent developments of statistical methods. The idea of a Budget for the purposes of National Finance has been familiar for some hundreds of years. This, as is well known, consists of a forecast of revenue balanced by a forecast of expenditure. It was not until comparatively recently that industry realized the benefits of such a system for business purposes. Since this time, however, the principle has been very widely adopted in America, and is looked upon as an essential part of any properly organized business. In England the value of budgetary control is becoming more generally appreciated, and the number of organizations applying the method is continually increasing.

This application of statistical methods involves a completely new aspect of the value of statistics. It has usually been the practice to compare data for the current year with data for the previous year or corresponding previous period. Budgetary control implies a new conception of the use of statistics, and consists of compiling a forecast of all data for the year, with which the actual figures are compared immediately they are available, instead of comparing them with similar figures for a previous period. The basis of budgetary control, therefore, is, at the commencement of each year, to make a scientific forecast of the figures of sales, purchases, wages and costs of production for the ensuing period, based upon the statistics of previous periods and the technical knowledge available. This then constitutes the budget, with which the actual figures have to be compared, so that this method of control enables comparisons to be made between the figures of the actual present, and predicted figures for the present, and the whole direction of policy is formulated on a preconceived Balance Sheet for the period under review.

The Sales Budget

The basis of budgetary control must of necessity rest upon a forecast of sales, and in this respect it should be realized that mere estimates, which are liable to be optimistic, are of no value. The data must be constructed from a scientific study of all the available information. The sales for previous years must be analysed to show secular and seasonal variations. The secular variations consist of the normal percentage increase or decrease per annum which can be found by a study of a moving average of the annual sales, thus eliminating exceptional circumstances. Thus if the secular trend shows a normal annual increase in sales of 2 per cent, it may be assumed primarily that for the ensuing year a similar increase will take place. In addition to this, however, general knowledge of other economic data must be taken into consideration. Market conditions and the trend of business activity in general must be carefully studied with the aid of statistical information published in the Press, Government and similar reports, and such aids as the London and Cambridge Economic Service. In many cases independent Market Research must be carried out before reliable forecasts can be made.

Since it is essential to check the actual sales with the predicted sales, it is not sufficient to make a forecast for the year as a whole. When this has been accomplished the figures must be allocated month by month, and to do this consideration must be given to the seasonal fluctuations which normally are found to take place. The extent of such fluctuations may be ascertained by comparing the actual monthly sales with the average monthly sales over the year, and the sales forecast analysed accordingly.

The sales figures thus obtained must then be analysed between the various classes of goods produced and the various salesmen, in order that quotas can be fixed, showing the volume of sales to be obtained each month from every representative.

In this respect it is a frequent practice to seek the suggestions of salesmen as to what they anticipate they can achieve in the period for which the forecast is made, and to take these factors into consideration in fixing the budget for sales.

Cost and Expenses Budgets

It will be realized that although the sales position is the most important feature of the method, the object of the organization is not to achieve a certain volume of turnover, but to make profits. It follows, therefore, that the sales budget must be followed by the budget of costs and the budget of administrative expenses.

A certain margin of profit is essential, and of course it is desirable that this margin should increase, so that if the profits are predicted as a certain ratio there is only a residue available to meet costs and administrative expenses. The budget of costs and the budget of administrative expenses are therefore built up to ensure a sufficient margin of profit, and are prepared in the utmost detail, based upon the experience of the past and the anticipated production required to meet the turnover which is forecast. Each department must be dealt with separately, forecasts being compiled to show the requirements in respect of wages, new equipment and materials, in the greatest possible detail. Financial requirements will be indicated by the proposed purchases and anticipated cash receipts from the forecast of sales. In this way provision can be made for capital requirements, additional borrowings or repayments of short-term loans. The budget of administrative expenses must be extended to cover all items involved in selling expenses and office expenses, detailed analyses being made available for all classes of expenditure on Salaries, Stationery, Advertising, etc.

Finally the whole of the independent budgets must be co-ordinated into a comprehensive budget to cover the whole operation of the business, involving sales, production costs, and administrative expenses, so that in effect a forecast of the final Balance Sheet is compiled.

Should the forecast show an adverse trend, after taking into account all the factors affecting the operation of the business, immediate action can be taken to avoid or minimize difficulties indicated by making provision for intensive advertising to increase or maintain sales, or on the other hand steps can be taken to reduce expenditure wherever this is possible by improving methods of production or restricting expenses in other directions.

The Application of the Budget

When the budget has been finally decided upon, each department is notified of the facts with which it is concerned, and it is incumbent on each department to work to its budget. Salesmen must obtain their quota in order that the sales budget will materialize. Statistics of material consumption for each productive department, together with wages and other costs, must be obtained each month and compared with the budget figures. Similarly data of administrative expenditure must also be compared monthly. In any case where expenditure on material, wages, or any other item exceeds the figures of the budget, an explanation will be required, and special steps must be taken to compensate the excess by reducing expenditure in the ensuing months.

The Advantages of Budgetary Control

The preparation of a budget involves the careful consideration of past statistics and general and special barometers of the future condition of the particular industry. By working to such a forecast it is possible to control stocks more closely, in such a way that capital is not tied up in stock which is not likely to be used for long periods, while, on the other hand, since requirements are anticipated in advance, purchases can be made when the prices of materials are at their lowest point.

Capital arrangements can also be made with the least possible expense, since the precise period when the additional capital can be liquidated is known in advance. In many cases a more even operation of plant can be obtained, resulting in the reduction of many overhead costs, and a more efficient employment of labour. In this way wastage due to labour turnover, the training of employees for short periods when demand is greatest, with the accompanying increased labour cost per unit of production, can be reduced by anticipating the peak production.

This anticipation assists in meeting customers' requirements promptly, and aids salesmen in obtaining their quotas. Since they are also keen to obtain the quota required of them, they have a definite objective which materially assists the maintenance of turnover, although it must be admitted that the introduction of quotas is, in the first instance, frequently strongly opposed by sales representatives.

Against the advantages claimed for budgetary control must be offset the disadvantages of working to a more or less rigid forecast. Salesmen are frequently afraid to obtain sales which are greater than their quota owing to the fact that they fear, if they increase their sales, future quotas will be increased until an impossibly high figure is reached. This disadvantage can usually be met by reasonable consideration of the sales which are possible to any representative.

On the production and administrative sides, there is again a possible disadvantage in too great a rigidity. No complaints will usually be made providing expenses and costs are within the budgetary figures. It is then probable that departmental heads will tend to be satisfied providing they can achieve this result, and in so doing fail to give sufficient attention to savings in production and administration, which would result in the expenses ultimately being much below the budget figures.

In practice, however, it is found that when the system is properly organized, its application is satisfactory and results in considerable advantages and increased efficiency. It must be emphasized that estimates are not sufficient, and any figures employed in budgetary control must be scientifically produced on the basis of sound statistical methods. In many cases this necessitates the establishment of a particular department for the purpose of compiling the budget and comparing the actual data with the predicted figures.

EXERCISES

1. What are the objects of budgetary control?
2. What procedure is necessary to assure proper budgetary control?

CHAPTER XIII

PUBLISHED STATISTICS

IN the conduct of business, statistics are employed in the solution of individual problems and the methods employed in such investigations are, of course, dependent upon the nature of the particular question under review. In addition, however, to the details compiled for such specific requirements of individual concerns, a vast amount of data is available in published form, designed to meet needs which are more or less general. It is only possible to survey briefly the most important classes of statistical data which are available for such general use, and to indicate the commercial uses of the data published.

The Population Census

Probably the best known series of official data are the details published as a result of the population census. The census has been taken every ten years since 1801, and, therefore, a vast amount of data is available for social investigations of many kinds. As already seen, the census is carried out by means of questionnaires forwarded to every dwelling place, the principal questions to which answers are demanded referring to age, sex, civil condition (whether married or single), number of children, occupation, and birthplace. The ultimate results of the investigation are, of course, affected by a great many circumstances over which no control is possible. For example, on the question regarding age, it is found that even under the pledge of secrecy many women understate their age, and, on the other hand, older people frequently have a tendency to overstate their age. Again, in classification of the data obtained, alterations are made from time to time in the allocation of administrative areas, which in some cases complicates comparisons, and makes a detailed study of the Registrar-General's report necessary for some purposes.

It is also an extremely difficult matter to classify the inhabitants of the country under a reasonable number of satisfactory occupational headings. In spite of these difficulties, however, data are compiled which are of great value for most general purposes. The final details are published under a number of classifications, covering the populations of Administrative Counties, County Boroughs, Urban and Rural Districts, separate volumes being published to deal with occupations, industries, and work places.

The results obtained are extensively used in the study of social problems, dealing with the increase or decrease in population of various centres, the effect of immigration and emigration, and such questions as overcrowding. Apart from the value of the records for such purposes, they are useful for many commercial purposes, and one or two examples of the manner in which they may be applied may be considered.

In establishing a manufacturing undertaking, great care must be exercised in the choice of a suitable site. Consideration will be given to the proximity of sources of supply of raw material, transport facilities, nearness to seaports and similar problems. It is also essential to ensure that in the locality decided upon, sufficient labour of the required skill will be available. Careful reference to the data contained in the population census will enable a decision to be made in this respect with reasonable accuracy.

In the marketing of finished products, care must be taken to allocate sales areas to the maximum advantage. Great assistance can be derived from the census figures in this problem, since the data will indicate the heavily populated areas which will require several representatives to carry on the work efficiently, as compared with those areas which, being sparsely populated although extensive, can be satisfactorily dealt with by one man.

It is the practice of many concerns to establish sales quotas representing preconceived estimates of the sales which should result from the work of individual representatives. In order to fix such quotas, a large amount of information is frequently required. As an example of the utility of the census figures in this direction, the method adopted by a certain jam manufacturing concern may be cited. The population for each sales area was estimated from the figures contained in the census. The average estimated consumption of jam per head of the population was then ascertained, and from this data the potential demand for jam was computed. On the basis of these figures quotas for each area were decided upon.

Trade

It is natural that a vast amount of data should be available relating to trade, industry, and commerce, and this information is available in many forms and for widely differing purposes. So far as international trade is concerned, monthly figures are published for the United Kingdom showing the quantity and value of imports and exports of various commodities, together with the principal sources or destinations, while the Annual Statement gives details of commodities imported and re-exported to different countries over a period of five years. The Statistical Abstract summarizes statistics of trade and shipping, and gives a considerable amount of detail with regard to individual commodities. The value of such statistical records as the above lies in the indications given concerning the progress made by trade in general.

More specific information is frequently required by business men to show the progress of individual industries, such data being necessary to form the basis of an opinion as to the general trend of trade. This information is available from a variety of sources, and includes such data as the following—

1. Monthly figures of the production of coal, pig iron, and steel for the United Kingdom and competitive foreign countries.
2. Monthly indices of security values.
3. Railway and shipping statistics showing the volume of various important commodities passing from month to month, and also the periodic monetary receipts of transport undertakings.
4. Consumption of iron, cotton, rubber, etc.

It will be seen that from a study and analysis of such comprehensive data as is available, it is possible to form a sound opinion of the present trading conditions, and also to make an estimate of the conditions likely to be experienced in the immediate future. Decisions of all kinds are also materially assisted by the market reports for separate trades which are published periodically.

Financial factors are adequately covered by statistical analyses which are regularly published, and include Discount Rates, Foreign Exchange Rates, and the volume of Bank Clearings.

Prices

Reference has already been made to the importance of following the trend of price movements. Every concern is closely affected by the price movements of raw materials essential for the manufacture of its products. Detailed information of the day-to-day price changes are provided in a large number of trade publications. Changes in general price movements are also available from the Statistical Abstract. It will be appreciated that, owing to the large variations in quality and other characteristics of the same commodity, difficulty is experienced in precisely defining the commodity to be measured, so that many factors must be taken into account in considering such statistical data as is available in this form. Standardization has, however, done much to improve the position, rendering it possible to obtain a reasonable average price by dividing the total value of a particular commodity imported by the quantity involved for the period under review. It must be borne in mind that such a figure represents an average price only, which is of use in considering the periodical fluctuations, but does not represent any definite market quotation. In addition, any comparisons are materially affected by variations in quality which must be taken into account in so far as information in this respect is available.

The above briefly indicates the data available in respect of price movements of individual commodities. In many investigations it is necessary to consider the trend of prices as a whole, and for this purpose index numbers are compiled by the Board of Trade, *Economist, Statist,* and other sources. The methods employed in obtaining these statistics and their commercial importance have already been considered in Chapter VII dealing with Index Numbers.

To obtain statistics of business activity as a whole is a very difficult problem. The *Economist* has, however, published an Index designed to indicate the trend of business activity by amalgamating the important factors in its composition. Such an index, although subject to limitations, is likely to prove a very valuable guide in the conduct of business affairs. Its main object is to show the progress or retrogression made by commerce as a whole, in order to indicate the possibility of improvement or recession in trade, which of course is precisely the information required by an industrialist in framing his commercial policy.

Additional information is available on this question as a result of the official Census of Production. In the United States of America such a census has regularly been made, but the first census of this kind was carried out in the United Kingdom in 1907. Questionnaires were then forwarded to manufacturers with a request for information to be given voluntarily, and from the replies received data was compiled which was considered to be representative of the whole. In the 1930 census of production, the data was made more complete by compulsorily requiring the information from all firms employing more than ten workers. The principal statistics made available by this census include for individual industries the selling value of the goods produced for sale, although not necessarily sold, and the amount paid out for materials, fuel, power and light, etc. The residue, known as the "net output," represents the amount available to meet wages, administrative and standing charges and profit on capital. The tables compiled are illustrated by Table XXXII.

TABLE XXXII

TABLE OF PRODUCTION

	1924		1930	
	Quantity 000 omitted	Value 000 omitted	Quantity 000 omitted	Value 000 omitted
		£		£
Articles produced . .	8,000	16,000	9,000	18,000
Repairs and Miscellaneous .		2,000		3,000
		18,000		21,000
Materials, etc. . . .		10,000		13,000
Net output . . .		8,000		8,000
Persons employed: Male . .	50,000		60,000	
Female .	30,000		40,000	
Net output per head . .	£100		£80	

In comparing the results obtained from various censuses, it must be remembered that the value of the goods produced is materially influenced by the relation of ruling prices at the various periods. Correct comparisons can only be made by reference to index numbers of prices. Extremely valuable information is, however, given of the progress of separate industries, so that it is possible to ascertain those industries in which progress has been made and those which have shown retrogression. The figures reveal the number of persons employed in the various industries and the tendency for transition to take place between the newer industries such as the electrical, silk and motor trades, which tend to employ more labour, and the older industries, such as coal and cotton, which show a decline in the numbers employed. Such statistics enable estimates to be made of the degree to which particular industries are affected by depression, as regards both the extent of employment and the magnitude of production.

Employment statistics, which are likewise available, are valuable from the commercial point of view and also as indicating the extent of unemployment. Data are published periodically in the Ministry of Labour *Gazette*. The main sources of information for these statistics are the records of trade unions and the returns made by the Employment Exchanges. The great majority of workers are covered by the National Unemployment Insurance Scheme. Although every description of worker is not covered, fairly comprehensive statistics are available. In respect of insured persons, however, the number actively employed is obtained every year when the insurance cards are sent in to the Employment Exchanges for the exchange of books, and the records are kept up to date by recording books issued to new entrants into industry and books returned in respect of deceased or retired workers. The extent of unemployment is obtained from the Employment Exchanges, who forward periodic returns to the Ministry of Labour. It is essential for every unemployed worker to deposit his insurance card with the exchange in order to obtain benefit under the Insurance Acts. The number unemployed at each centre is, therefore, available for the compilation of the necessary returns, so that in this manner statistics can be prepared to show the number of workers unemployed. The figures obtained are of value in that they indicate the extent to which industry is absorbing additional labour or is suffering from depression. The effect of seasonal fluctuations, which are important in normal times in such trades as the building trade and similar industries, is also shown.

Conclusion

In conclusion it will be seen that a great amount of statistical data is available for the guidance of the business man in regard to trade conditions in general. Assistance in formulating a conception of prevailing conditions may also be found in the summarized data published in periodic surveys of the state of trade, such as the monthly figures published by the *Economist* and other journals. Similar statistics are provided by the London and Cambridge Economic Service, which publishes monthly a bulletin of the General Business Position, together with indications of the possible tendencies for the immediate future. By having at their disposal such statistical data as is published in the forms outlined above, combined with adequate individual statistics, the responsible officials of a company can apply statistical methods with the maximum advantage to the conduct of the business, thereby materially assisting in its successful operation.

ANSWERS TO EXERCISES

CHAPTER I

1. The difference between statistics and mere figures lies largely in the fact that figures may be available representing a large and varied number of facts, but at the same time they do not assist in solving any important problem, neither can they be used for the purpose of making comparisons. Statistics, on the other hand, are scientifically compiled, after a clear conception of the problem has been formulated. Every figure obtained will, therefore, be relevant to some question under review, and will be compiled in such a manner that comparisons can be made and conclusions drawn, which, since they are based on facts numerically represented, can be relied upon as being accurate. They can thus be made the foundation for important decisions. On the contrary, it is obvious that mere figures, collected together without reference to any particular problem, cannot be so employed. They cannot be used to the maximum advantage, or as a guide to the solution of any problem, unless they satisfy all the requirements of statistical data.

2. It is frequently stated that statistics can be made to prove anything. This erroneous belief arises from the fact that in many cases arguments to support a fallacious statement are based upon figures treated in an unscientific manner, and wrongly called statistics. No consideration is given to the presentation of the figures to ensure that they satisfy the requirements necessary before being considered as satisfactory for statistical purposes. Thus it is frequent for conclusions to be made from figures which are incomplete, in that the whole of the information is not available, or that they are not sufficiently representative of the facts they are intended to present.

In addition to this, personal bias is allowed to enter into the consideration or calculation of the figures.

If, however, sufficient care is given to the facts presented, no fallacious statement may be supported. In accepting any argument based on figures, the accuracy of the figures should be verified and consideration given to the extent to which the data can be considered homogeneous, representative, complete, stable, and free from personal bias. Any data which satisfactorily comply with the rules governing the compilation of accurate statistics can then be relied upon for the basis of accurate conclusions.

CHAPTER II

1. The accuracy of sampling can only be absolutely verified by comparing the results of sampling with the results obtained from a consideration of the whole of the facts. This, however, is not practicable, since the object of employing the method of sampling lies in the fact that it obviates the necessity for dealing with the whole of the items which enter into an inquiry, in cases where it is either impossible or too costly to carry out a complete investigation. The accuracy may, however, be tested, and it is essential to do this in order to ensure that the samples taken are representative of the whole. In the first place, it should be ascertained that the samples have been taken in such a manner that no preference could have been given to any particular type. This may be ensured by choosing samples by lot, or, for example, by selecting every tenth or twentieth item.

To test the accuracy of sampling, several sets of samples should be taken. It is obvious, therefore, that if the method adopted gives a result which is truly representative of the whole, the results obtained from various sets of samples will approximately agree. This agreement should be obtained before the results of sampling are accepted as representative of the whole.

2. QUESTIONNAIRE

Name of Householder..

Age.............................

State whether employed or unemployed ....

Occupation ....

....

Wife of Householder ....

Age ....

Whether employed in any capacity...

If so, state nature of employment...

CHILDREN OVER 14 YEARS OF AGE, OR OTHER PERSONS OVER 14 YEARS OF AGE, RESIDENT WITH HOUSEHOLDER

Name	Sex	Age	Relationship	Occupation	Whether Employed or Unemployed

CHILDREN UNDER 14 YEARS OF AGE LIVING WITH THE HOUSEHOLDER

Name	Relationship	Age	Sex

3. (*a*) The allocation of expenditure of working-class budgets would best be investigated by forwarding questionnaires to representative families and tabulating the results. It is obviously impossible to investigate every possible case, and therefore the method of sampling must be employed.

(*b*) In carrying out an investigation concerning the extent to which acetylene welding is used, the questionnaire would again be employed and the results tabulated. The questionnaires would be sent to all those known to be users of such plant, and in this way a sufficiently representative sample would be obtained.

(*c*) In order to ascertain the effect of working overtime in a factory, the output for the factory analysed between the various departments would be carefully recorded. Similar figures would be recorded for the output of each shift, and the total output obtained compared with the output resulting from ordinary overtime. In each case the average output per worker per hour should be calculated to ascertain the effect on individual output of the altered conditions. Records of costs should also be taken, and a comparison made between the average labour cost per unit produced under the two arrangements.

CHAPTER III

1. (*a*) This is an unbiased error, and will be estimated as follows—

Average error for one item is 250, therefore total error is likely to be, $250 \times \sqrt{100} = 2{,}500.$

Expressed relatively, $\dfrac{2{,}500}{1{,}955{,}000} = {\cdot}00127$, or ·127%.

(*b*) This is a biased error estimated as follows—

Average error per item, 500.

Total error, $500 \times 100 = 50{,}000.$

Expressed relatively, $\dfrac{50{,}000}{1{,}955{,}000} = {\cdot}02557$, or 2·557%.

2. Column 3 suggests that the approximation should be made by approximating to the nearest 1,000, and will therefore result in an unbiassed error.

Column 4 will result in an approximation to the next lower 1,000, so that the error will be biased.

Column 5 means that the figures will be expressed in as few figures as possible, and it will be noted that this method enables the same conclusions to be made as any of the alternative methods of approximation suggested, since there is a large difference between the magnitude of the figures, and the obvious growth in magnitude is equally apparent—

Col. (1) Year	Col. (2) Actual	Col. (3) To nearest 1,000	Col. (4) 000's Omitted	Col. (5) Round Numbers
1	4,251,639	4,252	4,251	4·3 million
2	7,452,431	7,452	7,452	7·5 ,,
3	10,542,621	10,543	10,542	10·5 ,,
4	15,453,200	15,453	15,453	15·5 ,,
5	14,400,500	14,400	14,400	14·4 ,,
6	18,427,520	18,428	18,427	18·4 ,,

CHAPTER IV

1. Average income, $\frac{31{,}050}{135} = £230.$

Average tax, $\frac{1{,}552.10.0}{31{,}050} =$ 1s. in £ per head.

This average would be misleading if any very high incomes were included in the data. Such a case would exist if among the 135 inhabitants one person had an income of £10,000 per annum, so that this item alone would cause the average to be much higher than would otherwise be the case.

2. It is essential to calculate the weighted average thus—

$$\frac{(4{,}376{,}000 \times 36{\cdot}0) + (5{,}943{,}000 \times 34{\cdot}6) + (7{,}932{,}000 \times 35{\cdot}3)}{4{,}376{,}000 + 5{,}943{,}000 + 7{,}932{,}000}$$

$$= 35{\cdot}240 \text{ bushels.}$$

Using approximated weights the calculation would be—

$$\frac{(44 \times 36{\cdot}0) + (59 \times 34{\cdot}6) + (79 \times 35{\cdot}3)}{44 + 59 + 79} = 35{\cdot}242.$$

3. The average price of books sold is as follows—

$$\frac{(24 \times 3\text{d.}) + (25 \times 9\text{d}) + (80 \times 1/3) + (170 \times 1/9) + (155 \times 2/3) + (141 \times 2/9) + (92 \times 3/3) + (110 \times 3/9) + (80 \times 4/9) + (30 \times 5/9)}{24 + 25 + 80 + 170 + 155 + 141 + 92 + 110 + 80 + 30}$$

$$= \frac{121.2.9}{907} = 2/8{\cdot}05.$$

CHAPTER V

1. The value of the moving average lies in the fact that its application enables the underlying trend or development of a series to be studied. In considering the majority of important factors, whether relative to the turnover, profits or expenses of an individual business or the price of a commodity, it is found that from one period to another fluctuations take place. These short-term fluctuations are the result of a combination of factors which are exceptional. Thus the profits of a company may be very low in one year and show a considerable fall as compared with previous years. This fall, however, may be due to an expansion of the business to such a degree that the available capital was not sufficient. It is obvious from such an illustration that the fall in profits in no way reflected a decline in the development of the concern. The moving average

is, therefore, designed to eliminate as far as possible the effect of such fluctuations, in order to show the underlying tendency from which they represent incidental deviations.

The extent of short-term fluctuations is important and requires consideration, since their magnitude indicates the degree to which the matters under consideration are subject to fluctuation as a result of the operation of special factors. They may best be studied by calculating the difference between the moving average and the actual figures, and then plotting the deviations on a graph. In this manner it is possible to show graphically the importance of the fluctuations which have taken place.

Year	Profits	Five-year Moving Average
	£	£
1	10,000	—
2	8,000	—
3	7,000	12,400
4	12,000	13,600
5	25,000	16,800
6	16,000	19,000
7	24,000	23,600
8	18,000	24,600
9	35,000	25,400
10	30,000	25,600
11	20,000	30,000
12	25,000	33,000
13	40,000	33,000
14	50,000	36,000
15	30,000	33,000
16	35,000	—
17	10,000	—

The first five moving averages indicate a rapid growth of the business, followed by a period in which little progress was made. This was again followed by a period of improvement, which, however, does not appear likely to be maintained owing to the fall shown for the last period.

CHAPTER VI

1. The average is calculated as follows—

$$\frac{\begin{array}{c}(650 \times 14\text{s.}) + (432 \times 17\text{s.}) + (500 \times 18\text{s.}) + (271 \times 19\text{s.}) \\ + (250 \times 20\text{s.}) + (190 \times 21\text{s.}) + (184 \times 22\text{s.}) + (179 \times 23\text{s.}) \\ + (165 \times 24\text{s.}) + (55 \times 25\text{s.}) + (65 \times 26\text{s.}) + (102 \times 27\text{s.}) \\ + (73 \times 28\text{s.})\end{array}}{\begin{array}{c}650 + 432 + 500 + 271 + 250 + 190 + 184 + 179 + 165 \\ + 55 + 65 + 102 + 73\end{array}}$$

$$= \frac{2{,}978.11.0}{3{,}116} = 19/1{\cdot}41.$$

The data refer to a total of 3,116.

The median lies, therefore, between the 1,558th and 1,559th items.

Both these items lie within the third group, so that the median can be stated to be 18s. per week.

2. A consideration of the data presented leads to the following conclusions—

Machine No. 1. This machine shows a steady output which fluctuates only within a small margin of the modal output. It may, therefore, be assumed that the machine is being operated efficiently.

Machine No. 2. This machine falls far short of the modal output. Unless the age or condition of the machine is such as to be an explanation of the low output, the figures suggest that the maximum output is not being obtained, and further investigations are called for.

Machine No. 3. The output from this machine is invariably greater than the mode. This may be due to the greater efficiency of the operator, but care should be taken that the output is not being obtained at the expense of the life of the machine.

Machine No. 4. The output from this machine fluctuates considerably, and the cause should be ascertained. This may be due to breakdowns causing the output to fall for some days, or may be the result of erratic work on the part of an operator, working efficiently on some days but not so on others.

CHAPTER VII

1. CALCULATION OF INDEX NO.

(1) Commodity	(2) Weight	(3) 1925 Price as Percentage of 1914	(4) Product of (2) and (3)
Beef, *British*—			
Ribs . . .	12	180·00	2,160·00
Flank . . .	12	166·66	1,999·92
Beef, *Chilled* or *Frozen*—			
Ribs . . .	12	157·14	1,885·68
Flank . . .	12	120·00	1,440·00
Mutton, *British*—			
Legs . . .	6	200·00	1,200·00
Breast . . .	6	200·00	1,200·00
Mutton, *Chilled* or *Frozen*			
Legs . . .	6	171·43	1,028·58
Breast . . .	6	150·00	900·00
Bacon	19	163·64	3,109·16
Ham	20	200·00	4,000·00
Bread	50	183·33	9,166·50
Tea	22	166·66	3,666·52
Sugar	19	200·00	3,800·00
Milk	25	200·00	5,000·00
Butter . . .	20	185·71	3,714·20
Cheese . . .	10	155·55	1,555·50
Margarine . . .	10	100·00	1,000·00
Eggs	19	200·00	3,800·00
Potatoes . . .	18	222·22	3,999·96
Total . .	304		54,626·02

$$\text{Index No.} = \frac{54{,}626\cdot 02}{304} = 179\cdot 69$$

Share	1913 = 100	1922 Prices as Percentage of 1913	1923 Prices as Percentage of 1913	1930 Prices as Percentage of 1913
X and Co., Ltd.	100·00	76·66	108·33	106·66
Associated Engineers, Ltd.	100·00	142·86	157·14	107·14
Coals, Ltd.	100·00	252·63	305·26	252·63
Electricity, Ltd.	100·00	143·75	125·00	140·62
Machinery, Ltd.	100·00	45·00	50·00	35·00
	500·00	660·90	745·73	642·05
Index	100·00	132·18	149·15	128·41

As pointed out in the exercise, many more shares would require to be included in order to give a comprehensive index number. It would also be necessary to give some consideration to weights which should be employed.

CHAPTER VIII

1. Skeleton table to show earnings of men in the cotton trade—

Town	Under 20s.	20s. and under 25s.	25s. and under 30s.	30s. and under 35s.	35s. and over	Total Number	Average Earnings	
							s.	*d.*
1	9	67	29	4	—	109	23	–
2	1	9	20	—	—	30	24	10
3	10	214	75	—	—	299	23	3
4	37	406	361	23	6	833	24	10
5	69	1,293	1,669	121	14	3,166	35	5
6	20	190	127	6	6	349	24	6
7	185	1,448	1,942	402	86	4,063	25	11
8	88	606	203	33	19	949	24	4
9	258	756	416	72	12	1,514	23	4
10	122	535	147	6	1	811	22	7
Total	799	5,524	4,989	667	144	12,123	24	11
Percentage	6·5	45·6	41·2	5·5	1·2	100·00		

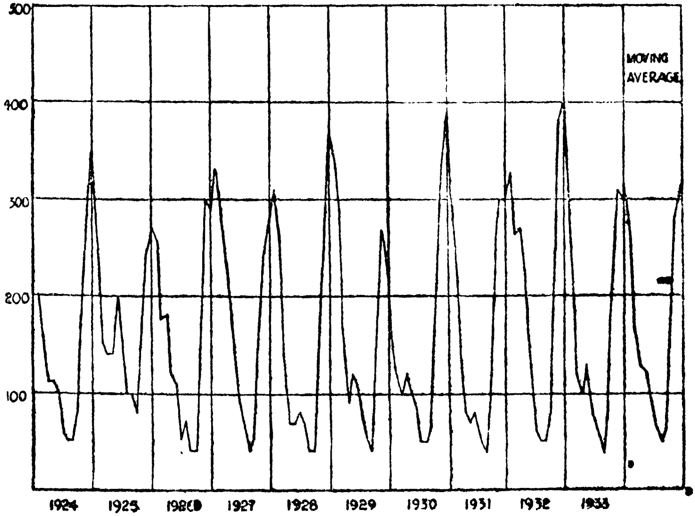
(*N.B.* Hypothetical figures have been included to show application of tables.)

2. Table to show numbers employed in various industries at 1911 and 1921 censuses—

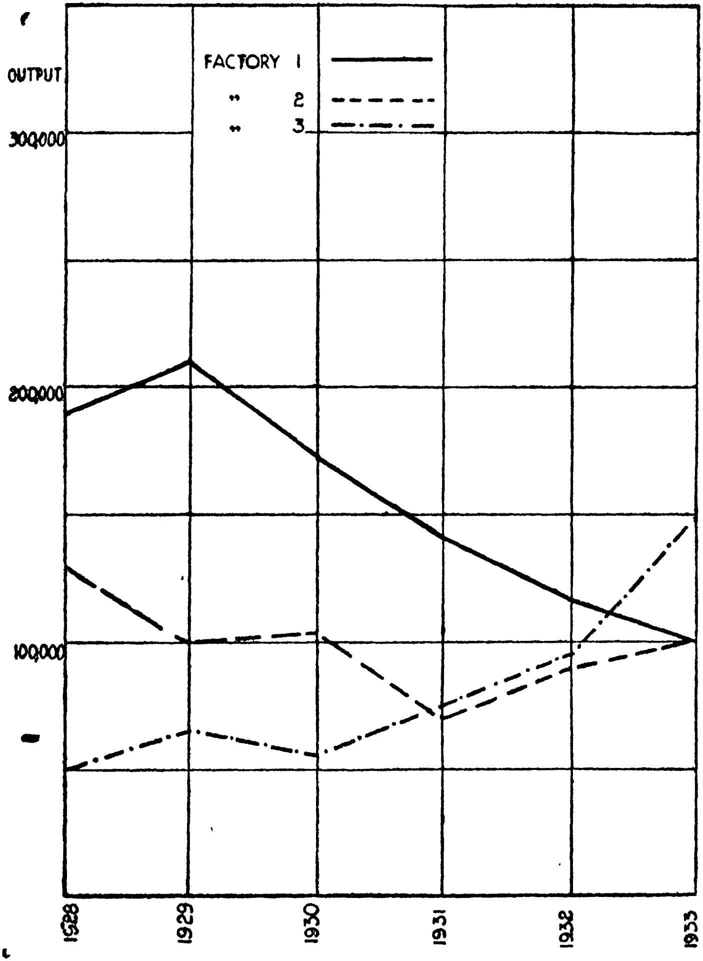
Age Group	BOOT AND SHOE		COAL		STEEL		ELECTRICAL	
	1911	1921	1911	1921	1911	1921	1911	1921
20 and under 25								
25 and under 30								
30 and under 35								
35 and under 40								
40 and under 45								
45 and under 50								
50 and over								
Total .								

CHAPTER IX

1.

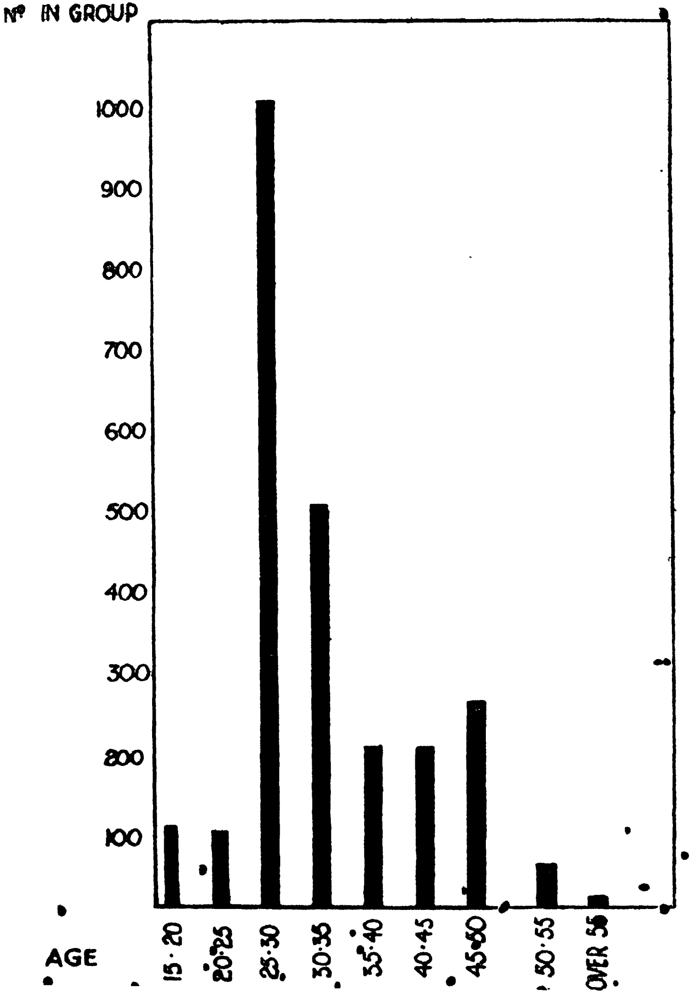


2. Graph showing comparison of output for Factories 1, 2, and 3 for the years 1928–1933.

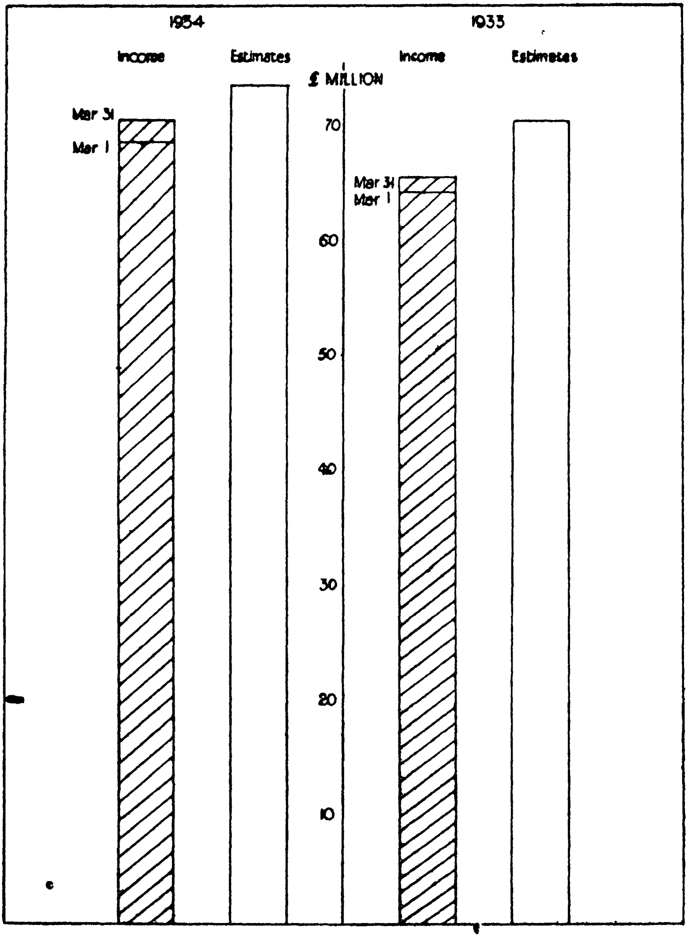


CHAPTER X

1. Diagram showing distribution according to age of the employees of a factory. (Bar Diagram.)



2. Diagram showing comparison between income tax collections, 1934 and 1933, and also comparison between actual receipts and estimates for the two years.



CHAPTER XI

1. Calculation of coefficient of Correlation.

RECEIPTS				EXPENSES			
Item	Pence per Mile	Deviation (x)	(x^2)	Pence per Mile	Deviation (y)	(y^2)	(xy)
1	9	—	—	6	—	—	—
2	9	—	—	7	+ 1	1	—
3	8	− 1	1	5	− 1	1	1
4	7	− 2	4	4	− 2	4	4
5	6	− 3	9	3	− 3	9	9
6	7	− 2	4	6	—	—	—
7	10	+ 1	1	7	+ 1	1	1
8	11	+ 2	4	7	+ 1	1	2
9	13	+ 4	16	8	+ 2	4	8
10	10	+ 1	1	7	+ 1	1	1
Total	90	16	40	60	12	22	+26

Standard Deviation Receipts $\sqrt{\frac{40}{10}} = 2{\cdot}00.$

,, ,, Expenses $\sqrt{\frac{22}{10}} = 1{\cdot}48.$

$$\text{Coefficient of Correlation} = \frac{26}{10 \times 2{\cdot}00 \times 1{\cdot}48} = \frac{26}{29{\cdot}6} = {\cdot}\dot{8}\dot{8}.$$

From the above it can be stated that the correlation is pronounced.

2. (*a*) Calculation of Average Deviation from Arithmetic Average.

Receipts				Expenses		
Item	Pence per Mile. (f)	Deviation (d)	Product ($f \times d$)	Pence per Mile (f_1)	Deviation (d_1)	Product ($f_1 \times d_1$)
1	9	—	—	6	—	—
2	9	—	—	7	1	7
3	8	1	8	5	1	5
4	7	2	14	4	2	8
5	6	3	18	3	3	9
6	7	2	14	6	—	—
7	10	1	10	7	1	7
8	11	2	22	7	1	7
9	13	4	52	8	2	16
10	10	1	10	7	1	7
Total	90	16	148	60	12	66

$$\text{Average Deviation Receipts } \frac{148}{90} = 1{\cdot}63.$$

$$\text{,, \quad ,, \quad Expenses } \frac{66}{60} = 1{\cdot}1.$$

2. (*b*) Calculation of Average Deviation from Median.

The median in each case lies between the 5th and 6th items after arranging in order of magnitude, and can therefore be stated to be 9d. in the series relating to receipts and 6·5d. in the series relating to expenses.

	RECEIPTS			EXPENSES		
Item	Pence per Mile (f)	Deviation (d)	Product ($f + d$)	Pence per Mile (f_1)	Deviation (d_1)	Product ($f_1 \times d_1$)
1	9	—	—	6	·5	3·0
2	9	—	—	7	·5	3·5
3	8	1	8	5	1·5	7·5
4	7	2	14	4	2·5	10·0
5	6	3	18	3	3·5	10·5
6	7	2	14	6	·5	3·0
7	10	1	10	7	·5	3·5
8	11	2	22	7	·5	3·5
9	13	4	52	8	1·5	12·0
10	10	1	10	7	·5	3·5
Total	90		148	60		60·0

Average Deviation from Median—

$$\text{Receipts } \frac{148}{90} = 1{\cdot}63.$$

$$\text{Expenses } \frac{60}{60} = 1{\cdot}00.$$

CHAPTER XII

1. The objects of Budgetary Control may be summarized as follows—

(*a*) To establish an accurate forecast of sales which will provide a basis for production and a foundation for practical and effective sales policy.

Subordinate to the main object of the sales budget, quotas are fixed as a guide and incentive to each individual sales representative, with a view to the achievement of the actual figures provided for in the budget.

(*b*) To provide a preconceived plan of manufacture so that materials can be purchased to the best advantage. This involves the purchase of materials during those periods when the prices tend to be at their lowest seasonal points, and also the careful control of stocks in such a way that capital is not unnecessarily bound up in stocks which are not likely to be used for long periods.

The planning of manufacture is also designed to eliminate wastage due to fluctuations in employment, and also to ensure that where production is seasonal supplies will be available to meet the peak demand, so that customers' requirements can be met without delay.

(*c*) To provide a financial guide to the capital requirements over the period under review. Receipts and expenses are forecast so that it is possible to see when the necessity to obtain additional capital by temporary borrowing will arise, the amount necessary, and the period over which it is required. This enables such capital to be raised in time to meet the demand for it, and at advantageous rates.

(*d*) Finally, to provide a complete plan on which the business can be controlled. On this basis it is possible to arrange selling policy and advertising schemes in such a manner that sales are obtained in accordance with the forecast, while on the other hand expenses can be kept within the budgetary figures, in order that the preconceived margin of profit may be obtained.

2. Proper budgetary control can only be achieved when the system is efficiently organized. It is not sufficient that mere estimates should be employed, and therefore the budget must be soundly constructed on statistical methods. Comprehensive data must be available as regards the past, in order to ascertain the probable trend of sales and expenditure. This information must be used as the basis for the computation of the budget, but current economic factors must also be taken into account to ascertain to what extent the business is likely to be influenced by general trade, monetary and labour conditions, whilst the prevailing state of the particular industry must also receive consideration.

The budget must be carefully analysed between the various departments, so that each departmental head may be made aware of the forecast in so far as it affects his responsibility. This enables each responsible official to see that expenditure provided for is not exceeded. The whole must be carefully co-ordinated with proper regard to the interdependence of the various departments.

Finally, it is vitally important that provision is made for comparing actual results with the budget periodically during the year. While the complete budget will relate to trading over a period of twelve months, the data must be allocated month by month and compared with the actual performance of each department and that of the business as a whole. This is important since, for example, it is useless to find at the end of nine months that expenditure has practically reached that budgeted for the year, and that sales have not materialized.

Thus when comparisons are made between the actual figures and the forecast month by month, it is possible to correct any variation over the ensuing period, and make any revisions necessary to the budget. As an illustration, sales may exceed those forecast, thus automatically increasing the actual expenses of production. Such a condition would necessitate the complete revision of the budget. This means that, although it is essential to work to the budget, a certain flexibility is essential to ensure that any savings in manufacturing costs, not provided for in the budget, may be taken advantage of, and that any tendency for sales to increase beyond expectations, may be properly provided for in the manufacturing and financial programmes.

INDEX

PAGE
ABSOLUTE error 34
Accuracy and approximation 26
Approximation 26, 27
——, error in 32
—— in compiling tables 100
——, methods of 30
——, purposes of 29
Arithmetic average 38 *et seq.*
——, advantages and disadvantages of 39, 54
——, progressive 43
Average, arithmetic 38 *et seq.*, 126
—— deviation 127
—— —— from median 130
——, moving 43, 56 *et seq.*
——, weighted 39, 44 *et seq.*

BALANCE of trade 29
Bar diagram 120
Biased Errors 32, 33
Block diagram 120
Budgetary control 138
—— ——, advantages of 142
—— ——, application of 142

CENSUS of population 145
—— of production 150
Coefficient of correlation 133
—— ——, calculation of 134
—— ——, interpretation of 135
—— ——, inverse 135
Coefficient of dispersion, average 128
—— ——, standard 130
Comparability 18, 19
Completeness 18, 22
Correlation 126 *et seq.*, 131
——, coefficient of 133
——, graphic presentation of 133
——, inverse 135
——, method of obtaining 132
Cost and expenses budget 140
—— of living index numbers 86

PAGE
DATA, characteristics of 18
——, compiling 8
——, presentation of 3, 92, 102, 117
——, sources of 9
Deviation 126
——, average 127
——, standard 127
Diagrams 117 *et seq.*
——, application of area 122
——, —— of circle 123
——, bar 120
——, block 120
——, pictogram 121

Economist index number 86
Errors 32
——, absolute 34
——, biased 32, 33
——, estimation of 34
——, measurement of 33 *et seq.*
——, relative 34
——, unbiassed 32

GRAPHIC method 102 *et seq.*
—— ——, logarithmic 110
—— ——, natural scale 110

HARVARD University Committee 114
Homogeneity 18, 20

INDEX numbers 78 *et seq.*, 150
—— ——, advantages and disadvantages of 84
—— ——, compilation of 80
—— ——, cost of living 86
—— ——, *Economist* 86
—— ——, published 85 *et seq.*
—— ——, *Statist* 86
—— ——, weighted and unweighted 82
—— ——, weights to be employed 85
—— of business activity 88
—— of national prosperity 88

LAW of statistical regularity 14
—— —— ——, practical application of 15
Logarithmic graphs 110
London School of Economics 114, 140

PAGE

MARKET research 17, 140
Measurement of error 33
Median 66 *et seq.*, 126
——, commercial value of 75
——, location of 71
Mode 66, 69 *et seq.*, 126
——, commercial value of 75
——, location of 71
Moving average 56
—— ——, application of 57 *et seq.*

PERCENTAGES, incorporation in tables 94, 99
Personal bias 18, 23
Pictograms 121
Population census 145
Presentation of data 3, 92, 102, 117
Prices, statistics of 85, 86, 149
Published statistics 85, 139, 145, 152

QUESTIONNAIRE 11
Quotas 108, 140

RAILWAY and shipping statistics 147
Range 128
Relative error 34

SALES budget 139
Sampling, application of 15
——, method of 13, 23
Sauerbeck's course of average prices 85
Short term fluctuations 59, 60
Stability 18, 23
Standard deviation 127, 129, 134
—— from median 130
Statist index numbers 86
Statistical regularity, law of 14
Statistics, application of 6
——, definition of 3
——, employment 152
——, financial 9
——, methods 3
——, objects of 2
——, production 10
——, published 85, 139, 145
——, trade 10, 147
——, wastage 10

PAGE
TABLES, approximation in 100
——, compilation of 94, 95, 96
——, incorporation of percentages in 94, 99
Tabulation . . . . 91 *et seq.*
——, function of 91
Trade statistics 147
Trading accounts, comparisons relating to 5
Trend 57, 60, 62, 63, 106

UNBIASED errors 32

WEIGHTED average . . . . 39, 44 *et seq.*
—— ——, application of 51
Weights 45
——, accuracy of 48, 49
——, importance of 48